Structure Formation independent of Cold Dark Matter

P. G. Miedema*
Netherlands Defense Academy
Hogeschoollaan 2,
NL-4818 CR Breda, The Netherlands

W. A. van Leeuwen
University of Amsterdam, Institute for Theoretical Physics,
Valckenierstraat 65,
NL-1018 XE Amsterdam, The Netherlands
(Dated: February 6, 2010)

It is shown that a first-order cosmological perturbation theory for Friedmann-Lemaître-Robertson-Walker universes admits one and only one gauge-invariant variable which describes the perturbation to the energy density and which becomes equal to the usual energy density of the Newtonian theory of gravity in the limit that all particle velocities are negligible with respect to the speed of light. The same holds true for the perturbation to the particle number density.

A cosmological perturbation theory based on these particular gauge-invariant quantities is more precise than any earlier first-order perturbation theory. In particular, it explains star formation in a satisfactory way, even in the absence of cold dark matter. In a baryon-only universe, the earliest stars, the so-called Population III stars, are found to have masses between 400 and 100,000 solar masses with a peak around 3400 solar masses. If cold dark matter is present then the star masses are between 130 and 13,000 solar masses with a peak around 450 solar masses. They come into existence between 100 Myr and 1000 Myr. At much later times, star formation is possible only in high density regions, for example within galaxies. Late time stars may have much smaller masses than early stars. The smallest stars that can be formed have masses of 0.2–0.8 solar mass, depending on the initial internal relative pressure perturbation.

It is demonstrated that the Newtonian theory of gravity cannot be used to study the evolution of cosmological density perturbations.

PACS numbers: 98.80.-k, 04.25.Nx, 98.65.Dx, 98.62.Ai, 98.80.Jk

Contents

*Electronic address: pieter.miedema@gmail.com

I. INTRODUCTION

It is well-known, or formulated more precisely, it is generally accepted, that in a universe filled with only 'ordinary matter,' i.e. elementary particles and photons but not *cold dark matter* (CDM), the linear perturbation theory predicts a too small growth rate to account for star formation in the universe. In this article we establish that this is not true. The reason brought forward in all former treatises on first-order cosmological density perturbations is that the growth of a relative density perturbation in the era after decoupling of radiation and matter, given by

$$\delta(t) = \delta(t_{\rm dec}) \left(\frac{t}{t_{\rm dec}}\right)^{2/3}, \quad t \geq t_{\rm dec}, \tag{1}$$

is insufficient for relative density perturbations as small as the observed value $\delta(t_{\rm dec}) \approx 10^{-5}$ to reach the non-linear phase for times $t \leq t_{\rm p}$, where $t_{\rm p} = 13.7\,\text{Gyr}$, the present age of the universe, and $t_{\rm dec} = 380\,\text{kyr}$, the time of decoupling of matter and radiation. This generally accepted conclusion is suggested by the solutions (E8), with $w = 0$, of the standard relativistic evolution equation (E6) for linear density perturbations in the era after decoupling of matter and radiation: this equation yields a growth rate which is much too low to allow for star formation within 13.7 Gyr. Therefore, researchers in the field of structure formation have to assume that a significant amount of CDM had contracted already before decoupling in order to explain in their simulations the formation of large-scale structure after decoupling [1, 2].

The purpose of this article is to show that the formation of structure can be explained whether or not CDM is present. Our treatise is independent of a particular system of reference and yields results which describe the evolution of small density perturbations in the radiation-dominated era and in the era after decoupling of matter and radiation. These remarkable and most satisfactory results are a direct consequence of two facts only. Firstly, we use gauge-invariant expressions for the first-order perturbations to the energy density, $\varepsilon^{\rm gi}_{(1)}$, and particle number density, $n^{\rm gi}_{(1)}$. Secondly, in the *dark ages* of the universe (i.e. the epoch between decoupling and the ignition of the first stars) a density perturbation from which stars will eventually be formed, has to cool down [3, 4] in its linear phase in order to grow. Consequently, the growth of density perturbations can only be described by a realistic equation of state for the pressure in combination with the combined First and Second Laws of thermodynamics (167). Therefore, we use an equation of state for the pressure of the form $p = p(n, \varepsilon)$, where n is the particle number density and ε is the energy density.

Our more sophisticated treatise, which in first-order perturbation theory is explicitly gauge-invariant, and which uses an equation of state of the form $p = p(n, \varepsilon)$ rather than of the form $p = p(\varepsilon)$, makes it possible to explain structure formation even in the absence of CDM.

II. MAIN RESULTS

In this article no assumptions or approximations (other than linearization of the Einstein equations and conservation laws) have been made in order to reach our conclusions. The only assumption we have made is that Einstein's General Theory of Relativity is the correct theory that describes gravitation in our universe on all scales and from the onset of the radiation-dominated era up to the present time.

In order to study structure formation in the universe, one needs the linearized Einstein equations. The derivation of the evolution equations (201) for relative density fluctuations in a Friedmann-Lemaître-Robertson-Walker (FLRW) universe filled with a perfect fluid with an equation of state $p = p(n, \varepsilon)$, is one of the main subjects of this article. Since we use a more general equation of state we are forced to derive all basic equations from scratch, instead of taking them from well-known textbooks or renowned articles. This has the advantage that our article is self-contained and that our results can easily be checked.

In Section X A we show, using the background (i.e. zeroth-order) Einstein equations and conservation laws connected with the combined First and Second Laws of thermodynamics, that the universe as a whole expands adiabatically. This is a well-known result. One of the new results of our treatise is that *local* density perturbations evolve *diabatically*. This has been made clear in Section X D. Only in the non-relativistic limit, where $\varepsilon = \varepsilon(n)$ and $p = 0$, local density perturbations evolve adiabatically.

In the literature about the subject, all efforts to construct a gauge-invariant cosmological perturbation theory yield a second-order differential equation for the density contrast function $\delta \equiv \varepsilon_{(1)}/\varepsilon_{(0)}$, with or without an entropy related source term. For that matter, our treatise is no exception, as (201a) demonstrates. In contrast to the perturbation theories developed in the literature, we find also a first-order differential equation, namely (201b). This equation results from the incorporation of the equation of state $p = p(n, \varepsilon)$ for the pressure and the particle number conservation law $(nu^\mu)_{;\mu} = 0$, (45). The consequences of equation (201b) are radical: this equation implies that density perturbations in the total energy density are gravitationally coupled to density perturbations in the particle number density. This is the case for ordinary matter as well as CDM throughout the history of the universe from the onset of the radiation-dominated era until the present. As a consequence, perturbations in CDM evolve gravitationally in exactly the same way as perturbations in ordinary matter do. The assumption that CDM would have clustered already before decoupling and thus would have formed seeds for baryon contraction after decoupling is, therefore, questionable. This conclusion has, on different grounds, also been reached by Nieuwenhuizen *et al.* [5]. This may rule out CDM as a means to facilitate structure formation in the universe.

A. Manifestly Gauge-invariant Perturbation Theory and its Non-Relativistic Limit

In order to solve the structure formation problem of cosmology, we first develop in Sections IV–XI a *manifestly* gauge-invariant perturbation theory, i.e. both the evolution equations and their solutions are independent of the choice of a system of reference. In Section IV we show that there exist two and only two *unique* and gauge-invariant first-order quantities $\varepsilon^{\text{gi}}_{(1)}$ and $n^{\text{gi}}_{(1)}$ for the perturbations to the energy density and the particle number density. The evolution equations for the contrast functions $\delta_\varepsilon \equiv \varepsilon^{\text{gi}}_{(1)}/\varepsilon_{(0)}$ and $\delta_n \equiv n^{\text{gi}}_{(1)}/n_{(0)}$ are given by (201). From their derivation it follows that these equations include the —background as well as first-order— G_{00}- and G_{0i}-constraint equations; the G_{ij}-evolution equations; the conservation laws for the energy density; for the particle number density and for the momentum; and, finally, the combined First and Second Laws of thermodynamics. Taking into account the applicability of equations (201) to the open closed and flat FLRW universes filled with a perfect fluid described by an equation of state $p = p(n, \varepsilon)$, they are rather simple.

In Section XII we show that in the non-relativistic limit $v/c \to 0$ the quantities $\varepsilon^{\text{gi}}_{(1)}$ and $n^{\text{gi}}_{(1)}$ survive, and coincide with the usual energy density and particle number density [see (229) and (230)], whereas their gauge dependent counterparts $\varepsilon_{(1)}$ and $n_{(1)}$ (which are also gauge dependent in the non-relativistic limit) disappear completely from the scene. Finally, we show that, in first-order, the *global* expansion of the universe is not affected by *local* perturbations in the energy and particle number densities.

B. Large-Scale Perturbations: Confirmation of the Standard Knowledge

In order to compare our treatise on cosmological density perturbations with the standard knowledge, we consider a flat FLRW universe in the radiation-dominated era and in the era after decoupling of matter and radiation. For *large-scale* perturbations these two cases have been thoroughly studied by a large number of researchers from 1946 up till now, using the full set of linearized Einstein equations and conservation laws. Consequently, our refinement of their work cannot be expected to give results that differ much from those of the standard theory. Indeed, we have found that for large-scale perturbations our manifestly gauge-invariant treatise corroborates the outcomes of the standard theory in the large-scale limit of both eras, with the exception that we do *not* find, of course, the non-physical gauge mode $\delta_{\text{gauge}} \propto t^{-1}$ which plagues the standard theory.

For example, our perturbation theory yields in the radiation-dominated era the well-known solutions (258) $\delta_\varepsilon \propto t$ and $\delta_\varepsilon \propto t^{1/2}$ [6–11] and in the era after decoupling of matter and radiation we get the well-known solutions (310)

$\delta_\varepsilon \propto t^{2/3}$ [6–11] and $\delta_\varepsilon \propto t^{-5/3}$ [12, 13]. A new result is, however, that the solutions (258) as well as (310) follow from *one* second-order differential equation, namely (245a) and (293) respectively.

C. Small-Scale Perturbations and Star Formation: New Results

The major difference between our treatise and the standard treatise on the subject lays in the evolution of small-scale density perturbations. For a radiation-dominated universe, the standard theory yields oscillating density perturbations (352) with a *decaying* amplitude. In contrast, our theory yields oscillating density perturbations with an *increasing* (260) amplitude. This difference is entirely due to the presence of the spurious gauge modes in the solutions of the standard equations, as we will explain in detail in Section XVI A.

After decoupling of matter and radiation at $z = 1091$, the results of our treatise and the results found in literature differ also considerably. Just as is done in previous researches, we take as equations of state for the energy density and the pressure $\varepsilon = nm_{\mathrm{H}}c^2 + \frac{3}{2}nk_{\mathrm{B}}T$ and $p = nk_{\mathrm{B}}T$, respectively, in the background as well as in the perturbed universe. Since $m_{\mathrm{H}}c^2 \gg k_{\mathrm{B}}T$ throughout the matter-dominated era after decoupling, it follows that one may neglect the pressure $nk_{\mathrm{B}}T$ and kinetic energy density $\frac{3}{2}nk_{\mathrm{B}}T$ with respect to the rest-mass energy density $nm_{\mathrm{H}}c^2$ in the unperturbed universe and that in the perturbed universe one has $\delta_\varepsilon \approx \delta_n$. Therefore, one takes in the literature $\delta_\varepsilon = \delta_n$ and solves the (homogeneous) second-order evolution equation for δ_ε. As we have shown in Section XIII C 3, this yields slowly growing density perturbations with $\delta_p = \delta_\varepsilon$, i.e. the relative pressure perturbation is equal to the relative energy density perturbation, and a vanishing relative *matter* temperature perturbation δ_T.

In contrast to the standard method, our perturbation theory yields next to the usual second-order evolution equation (293) for δ_ε also a first-order evolution equation (278) for the difference $\delta_n - \delta_\varepsilon$. Therefore, we need not take δ_n exactly equal to δ_ε, so that in our treatise we may have $\delta_p \neq \delta_\varepsilon$ and $\delta_T \neq 0$. As a consequence, our resulting second-order evolution equation (298) becomes inhomogeneous and the initial relative matter temperature perturbation $\delta_T(t_{\mathrm{dec}}, \boldsymbol{x})$ enters the source term. This proves to be crucial for star formation. Although in a linear perturbation theory $|\delta_T(t, \boldsymbol{x})| \leq 1$, this quantity is not constrained to be as small as $\delta_\varepsilon(t_{\mathrm{dec}}, \boldsymbol{x}) \approx \delta_n(t_{\mathrm{dec}}, \boldsymbol{x}) \approx 10^{-5}$, as is demanded by WMAP-observations [14–18]. Since the gas pressure $p = nk_{\mathrm{B}}T$ is very low, its relative perturbation $\delta_p \equiv p^{\mathrm{gi}}_{(1)}/p_{(0)}$ and, accordingly, the matter temperature perturbation $\delta_T(t_{\mathrm{dec}}, \boldsymbol{x})$ could be large. We have shown that just after decoupling at $z = 1091$ negative relative matter temperature perturbations as small as -0.5% yields massive stars within 13.7 Gyr. The very first stars, the so-called Population III stars [19–21], come into existence between 10^2 Myr and 10^3 Myr and have masses between $4 \times 10^2\,\mathrm{M}_\odot$ and $10^5\,\mathrm{M}_\odot$, with a peak around $3.4 \times 10^3\,\mathrm{M}_\odot$. Stars lighter than $3.4 \times 10^3\,\mathrm{M}_\odot$ come into existence at later times, because their internal gravity is weaker. On the other hand, stars heavier than $3.4 \times 10^3\,\mathrm{M}_\odot$ also develop later, since they do not cool down so fast due to their large scale. These conclusions, which are valid only in a universe filled with a baryonic fluid, are outlined in Figure 1. However, if CDM is present then the peaks in Figure 1 are at $4.5 \times 10^2\,\mathrm{M}_\odot$, whereas for *hot dark matter* (HDM) the peaks are found to be at $8.2 \times 10^3\,\mathrm{M}_\odot$. The relation between the particle mass and the mass of a star will be explained in Section XIV B. The peaks in Figure 1 can be considered as the relativistic counterparts of the classical *Jeans mass*.

At much later times, star formation is still possible, however the mass of the stars may be much smaller. For example, if star formation starts at $z = 1$ or later then the smallest stars that can be formed have masses of $0.2\,\mathrm{M}_\odot$–$0.8\,\mathrm{M}_\odot$, depending on the initial internal relative matter temperature perturbation. Also the initial density perturbations must be considerable at late times in order to make star formation feasible: for star formation starting at $z = 1$ one must have $0.7 \lesssim \delta_n \approx \delta_\varepsilon < 1$. This shows that late time star formation is possible only in high density regions within galaxies, but not in intergalactic space. These findings are summarized in Figures 2 and 3. In contrast to the relativistic theory developed in this article the standard Newtonian theory of linear perturbations, which does *not* follow from Einstein's gravitational theory, predicts, as can be seen from Figure 4, a lower limit for star formation of $1.7\,\mathrm{M}_\odot$, implying that our Sun could not exist at all! This failure of the standard Newtonian perturbation theory can be attributed to the gauge mode which is present in the solution (359) of the standard equation (357).

We conclude that density perturbations in ordinary matter can account for star formation. There is no need to make use of alternative gravitational theories nor of the inclusion of CDM: the Theory of General Relativity can be used to explain structure in our universe. The important conclusion must be that Einstein's gravitational theory not only describes the *global* characteristics of the universe, but is also *locally* successful.

III. LINEAR PERTURBATIONS IN THE GENERAL THEORY OF RELATIVITY

Perturbation theories for FLRW universes are, in general, constructed along the following lines. First, all quantities relevant to the problem are divided into two parts: 'a background part' and 'a perturbation part.' The background parts are chosen to satisfy the Einstein equations for an isotropic universe, i.e. one chooses for the background quantities

the FLRW-solution. Because of the homogeneity, the background quantities depend on the time coordinate t only. The perturbation parts are supposed to be small compared to their background counterparts, and to depend on the space-time coordinate $x = (ct, \boldsymbol{x})$. The background and perturbations are often referred to as 'zeroth-order' and 'first-order' quantities respectively and we will use this terminology also in this article. After substituting the sum of the zeroth-order and first-order parts of all relevant quantities into the Einstein equations, all terms that are products of two or more quantities of first-order are neglected. This procedure leads, by construction, to a set of *linear* differential equations for the quantities of first-order. The solution of this set of linear differential equations is then reduced to a standard problem of the theory of ordinary, linear differential equations.

A. History

The first systematic study of cosmological density perturbations is due to Lifshitz [22, 23] (1946) and Lifshitz and Khalatnikov [6] (1963). They considered small variations in the *metric tensor* to study density perturbations in the radiation-dominated ($p = \frac{1}{3}\varepsilon$) and matter-dominated ($p = 0$) universe. The use of metric tensor fluctuations makes their method vulnerable to spurious solutions, the so called gauge modes. Adams and Canuto (1975) [7] extended the work of Lifshitz to a more general equation of state $p = w\varepsilon$, where w is a constant. In 1966 Hawking [24] presented a perturbation theory which is explicitly coordinate-independent. Instead of using the perturbed metric tensor, he considered small variations in the *curvature* due to density perturbations. In 1976, Olson [8] corrected and further developed the work of Hawking. He defined the density perturbation relative to co-moving proper time. The advantage of Olson's method is that the gauge mode, still present in his solutions, can be readily identified since gauge modes yield, in his theory, a vanishing curvature perturbation. Bardeen [12] (1980), was one of the first who realized that one should work with variables which are themselves gauge-invariant. He used in his work two different definitions of gauge-invariant density perturbations, which, in the small-scale limit, coincide with the usual density perturbation which is gauge dependent outside the horizon. Bardeen assumed that a gauge dependent perturbation becomes gauge-invariant as soon as the perturbation becomes smaller than the horizon. In Section XII we show that this assumption is invalid. Kodama and Sasaki [25] (1984) elaborated and clarified the pioneering work of Bardeen. Ellis, Bruni and Hwang *et al.* [26, 27] (1989) and Ellis and van Elst [28] (1998) criticized the work of Hawking and Bardeen and gave an alternative and elegant representation of density fluctuations. Their method is both fully covariant and gauge-invariant. Although the standard equations (355) and (362) are, according to Einstein's General Theory of Relativity, closely related via (E6), equation (362) follows from the theory of Ellis *et al.*, but equation (355) cannot be derived from their method. Moreover, they did not take into account the perturbed constraint equations. Mukhanov, Feldman and Brandenberger, in their 1992 review article [13] entitled 'Theory of Cosmological Perturbations' mentioned or discussed more than 60 articles on the subject, and, thereupon, suggested their own approach to the problem. Their method is also discussed in the textbook of Mukhanov [29]. A disadvantage of their method is that their perturbation theory does not yield the usual Poisson equation in the non-relativistic limit $v/c \to 0$, implying that their gauge-invariant quantities cannot be linked to their Newtonian counterparts. For an overview of the literature we refer to Mukhanov *et al.* [13] and the summer course given by Bertschinger [30]. A recent and integral overview of the construction of cosmological perturbation theories for flat FLRW universes is given by Malik and Wands [31].

The fact that so many studies are devoted to a problem that is nothing but obtaining the solution of a set of ordinary, linear differential equations is due to the fact that there are several complicating factors, not regarding the mathematics involved, but with respect to the physical interpretation of the solutions. As yet there is no consensus about which solution is the best. In this article we will demonstrate that there is one and only one solution to the problem how to construct gauge-invariant quantities. This then enables us to solve the problem of structure formation.

B. Origin of the Interpretation Problem

At the very moment that one has divided a physical quantity into a zeroth-order and a first-order part, one introduces an ambiguity. Let us consider the energy density of the universe, $\varepsilon(x)$, and the particle number density of the universe, $n(x)$. The linearized Einstein equations contain as *known* functions the zeroth-order functions $\varepsilon_{(0)}(t)$ and $n_{(0)}(t)$, which describe the evolution of the background, i.e. they describe the evolution of the unperturbed universe and they obey the unperturbed Einstein equations, and as *unknown* functions the perturbations $\varepsilon_{(1)}(x)$ and $n_{(1)}(x)$. The latter are the solutions to be obtained from the linearized Einstein equations. The sub-indexes 0 and 1, which indicate the order, have been put between round brackets, in order to distinguish them from tensor indices. In all calculations, products of a zeroth-order and a first-order quantity are considered to be of first-order, and are retained, whereas products of first-order quantities are neglected.

The ambiguity is that the linearized Einstein equations do not fix the quantities $\varepsilon_{(1)}(x)$ and $n_{(1)}(x)$ uniquely. In fact, it turns out that next to any solution for $\varepsilon_{(1)}$ and $n_{(1)}$ of the linearized Einstein equations, there exist solutions of the form

$$\hat{\varepsilon}_{(1)}(x) = \varepsilon_{(1)}(x) + \psi(x)\partial_0\varepsilon_{(0)}(t), \tag{2a}$$

$$\hat{n}_{(1)}(x) = n_{(1)}(x) + \psi(x)\partial_0 n_{(0)}(t), \tag{2b}$$

which also satisfy the linearized Einstein equations. Here the symbol ∂_0 stands for the derivative with respect to $x^0 = ct$. The function $\psi(x)$ is an arbitrary but 'small' function of the space-time coordinate $x = (ct, \boldsymbol{x})$, i.e. we consider $\psi(x)$ to be of first-order. We will derive (2) at a later point in this article; here it is sufficient to note that the perturbations $\varepsilon_{(1)}$ and $n_{(1)}$ are fixed by the linearized Einstein equations up to terms that are proportional to an arbitrary, small function $\psi(x)$, usually called a gauge function in this context. Since a physical quantity, i.e. a directly measurable property of a system, may not depend on an arbitrary function, the quantities $\varepsilon_{(1)}$ and $n_{(1)}$ cannot be interpreted as the real physical values of the perturbations in the energy density or the particle number density. But if $\varepsilon_{(1)}$ and $n_{(1)}$ are not the physical perturbations, what *are* the real perturbations? This is the notorious 'gauge problem' encountered in any treatise on cosmological perturbations. Many different answers to this question can be found in the literature, none of which is completely satisfactory; a fact which explains the ongoing discussion on this subject. In this article we show that there is a definitive answer to the gauge problem of cosmology.

IV. GAUGE-INVARIANT QUANTITIES

In the existing literature on cosmological perturbations, one has attempted to solve the problem that corresponds to the gauge dependence of the perturbations $\varepsilon_{(1)}$ and $n_{(1)}$ (2) in two, essentially different, ways. The first way is to impose an extra condition on the gauge field $\psi(x)$ [30, 32–34]. Another way to get rid of the gauge field $\psi(x)$ is to choose linear combinations of the matter variables $\varepsilon_{(1)}$, $n_{(1)}$ and other gauge dependent variables to construct gauge-invariant quantities. The latter method is generally considered better than the one where one fixes a gauge, because it not only leads to quantities that are independent of an undetermined function, as should be the case for a physical quantity, but it also does not rely on any particular choice for the gauge function.

The newly constructed gauge-invariant quantities are then shown to obey a set of linear equations, not containing the gauge function $\psi(x)$ anymore. These equations follow, by elimination of the gauge dependent quantities in favor of the gauge-invariant ones, in a straightforward way from the usual linearized Einstein equations, which did contain $\psi(x)$. In this way, the theory is no longer plagued by the gauge freedom that is inherent to the original equations and their solutions: $\psi(x)$ has disappeared completely, as it should, not with brute force, but as a natural consequence of the definitions of the perturbations to the energy and the particle number densities. This method is elaborated by Bardeen [12] and Mukhanov *et al.* [13]. From these two treatises on linear perturbation theory, which differ significantly from each other, one is tempted to conclude that gauge-invariant quantities can be constructed in many different ways, and that there is no way to tell which of these theories describes the evolution of density perturbations correctly. This, however, is not the case, as we will show in this article.

We follow the method advocated by Mukhanov *et al.* and Bardeen, but with gauge-invariant quantities which differ substantially from those used by these researchers. In fact, we will show that there exist *unique* gauge-invariant quantities

$$\varepsilon^{\mathrm{gi}}_{(1)} \equiv \varepsilon_{(1)} - \frac{\partial_0\varepsilon_{(0)}}{\partial_0\theta_{(0)}}\theta_{(1)}, \tag{3a}$$

$$n^{\mathrm{gi}}_{(1)} \equiv n_{(1)} - \frac{\partial_0 n_{(0)}}{\partial_0\theta_{(0)}}\theta_{(1)}, \tag{3b}$$

that describe the perturbations to the energy density and particle number density. In these expressions $\theta_{(0)}$ and $\theta_{(1)}$ are the background and perturbation part of the covariant four-divergence $\theta = c^{-1}U^{\mu}{}_{;\mu}$ of the cosmological fluid velocity field $U^{\mu}(x)$.

In Section IV A, we will show that the quantities $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ do *not* change if we switch from the old coordinates x^{μ} to new coordinates $\hat{x}^{\mu}$ according to

$$\hat{x}^{\mu} = x^{\mu} - \xi^{\mu}(x), \tag{4}$$

where the $\xi^{\mu}(x)$ $(\mu = 0, 1, 2, 3)$ are four arbitrary functions, considered to be of first-order, of the old coordinates x^{μ},

i.e. (4) is an infinitesimal coordinate transformation, or gauge transformation. In other words, we will show that

$$\hat{\varepsilon}^{\mathrm{gi}}_{(1)}(x) = \varepsilon^{\mathrm{gi}}_{(1)}(x), \tag{5a}$$

$$\hat{n}^{\mathrm{gi}}_{(1)}(x) = n^{\mathrm{gi}}_{(1)}(x), \tag{5b}$$

i.e. the perturbations (3) are independent of $\xi^\mu(x)$, i.e. gauge-invariant.

Since the background quantities depend on time, but not on the spatial coordinates, it will turn out that only the zero component of the gauge functions $\xi^\mu(x)$ occurs in the transformation of the first-order gauge dependent variables. We will call it $\psi(x)$:

$$\psi(x) \equiv \xi^0(x). \tag{6}$$

In the perturbation theory, the gauge function $\psi(x)$ is to be treated as a first-order quantity, i.e. as a small (or 'infinitesimal') change of the coordinates.

As yet, the quantities (3) are new: they have never been used before. *The fact that these quantities are unique follows immediately from the linearized Einstein equations for scalar perturbations, and, therefore, cannot be chosen arbitrarily.* Using the quantities (3), our theory reduces to the usual Newtonian theory (229) and (230) in the limit that the spatial part of the cosmological fluid velocity four-vector U^μ is small compared to the velocity of light.

A. Construction of Gauge-invariant First-order Perturbations

We now proceed with the proof that $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ are gauge-invariant, i.e. invariant under the general infinitesimal coordinate transformation (4). To that end, we start by recalling the defining expression for the Lie derivative of an arbitrary tensor field $A^{\alpha\cdots\beta}{}_{\mu\cdots\nu}$ with respect to a vector field $\xi^\tau(x)$. It reads

$$\begin{aligned}(\mathcal{L}_\xi A)^{\alpha\cdots\beta}{}_{\mu\cdots\nu} = {} & A^{\alpha\cdots\beta}{}_{\mu\cdots\nu;\tau}\xi^\tau \\ & - A^{\tau\cdots\beta}{}_{\mu\cdots\nu}\xi^\alpha{}_{;\tau} - \cdots - A^{\alpha\cdots\tau}{}_{\mu\cdots\nu}\xi^\beta{}_{;\tau} \\ & + A^{\alpha\cdots\beta}{}_{\tau\cdots\nu}\xi^\tau{}_{;\mu} + \cdots + A^{\alpha\cdots\beta}{}_{\mu\cdots\tau}\xi^\tau{}_{;\nu},\end{aligned} \tag{7}$$

where the semi-colon denotes the covariant derivative. At the right-hand side, there is a term with a plus sign for each lower index and a term with a minus sign for each upper index. Recall also, that the covariant derivative in the expression for the Lie derivative may be replaced by an ordinary derivative, since the Lie derivative is, by definition, independent of the connection. This fact simplifies some of the calculations below.

Now, let $\{x^\mu\}$ and $\{\hat{x}^\mu = x^\mu - \xi^\mu(x)\}$ be two sets of coordinate systems, where $\xi^\mu(x)$ is an arbitrary —but infinitesimal, i.e. in this article, of first-order— vector field. Then the components $\hat{A}^{\alpha\cdots\beta}{}_{\mu\cdots\nu}(x)$ of the tensor A with respect to the new coordinates $\hat{x}^\mu$ can be related to the components of the tensor $A^{\alpha\cdots\beta}{}_{\mu\cdots\nu}(x)$, defined with respect to the old coordinates $\{x^\mu\}$ with the help of the Lie derivative. Up to and including terms containing first-order derivatives one has

$$\hat{A}^{\alpha\cdots\beta}{}_{\mu\cdots\nu}(x) = A^{\alpha\cdots\beta}{}_{\mu\cdots\nu}(x) + (\mathcal{L}_\xi A)^{\alpha\cdots\beta}{}_{\mu\cdots\nu}(x) + \cdots . \tag{8}$$

For a derivation of this expression, see Weinberg [35], Chapter 10, Section 9.

Note that x in the left-hand side corresponds to a point, P say, of space-time with coordinates x^μ in the coordinate frame $\{x\}$, while in the right-hand side x corresponds to another point, Q say, with exactly the same coordinates x^μ, but now with respect to the coordinate frame $\{\hat{x}\}$. Thus, (8) is an expression that relates one tensor field A at two different points of space-time, points that are related via the relation (4).

The following observation is crucial. Because of the general covariance of the Einstein equations, they are invariant under general coordinate transformations $x \to \hat{x}$ and, in particular, under coordinate transformations given by (4). Hence, if some tensorial quantity $A(x)$ of rank n ($n = 0, 1, \ldots$) satisfies the Einstein equations with as source term the energy-momentum tensor T, the quantity $\hat{A}(x) = A(x) + \mathcal{L}_\xi A(x)$ satisfies the Einstein equations with source term $\hat{T}(x) = T(x) + \mathcal{L}_\xi T(x)$, for a universe *with precisely the same physical content.* Because of the linearity of the linearized Einstein equations, a linear combination of any two solutions is also a solution. In particular, $\mathcal{L}_\xi A$, being the difference of A and $\hat{A}$, is a solution of the linearized Einstein equations with source term $\mathcal{L}_\xi T$. In first-order, $\mathcal{L}_\xi A(x)$ may be replaced by $\mathcal{L}_\xi A_{(0)}(t)$, where $A_{(0)}(t)$ is the solution for $A(t)$ of the zeroth-order Einstein equations. The freedom to add a term of the form $\mathcal{L}_\xi A_{(0)}(t)$, with ξ^μ ($\mu = 0, 1, 2, 3$) four arbitrary functions of first-order, to any solution of the Einstein equations of the first-order, is the reason that none of the first-order solutions is uniquely

defined, and, hence, does not correspond in a unique way to a measurable property of the universe. This is the notorious gauge problem. The additional terms $\mathcal{L}_\xi A_{(0)}(t)$ are called 'gauge modes.'

Combining (7) and (8) we have

$$\begin{aligned}\hat{A}^{\alpha\cdots\beta}{}_{\mu\cdots\nu}(x) &= A^{\alpha\cdots\beta}{}_{\mu\cdots\nu}(x) + A^{\alpha\cdots\beta}{}_{\mu\cdots\nu;\tau}\xi^\tau \\ &- A^{\tau\cdots\beta}{}_{\mu\cdots\nu}\xi^\alpha{}_{;\tau} - \cdots - A^{\alpha\cdots\tau}{}_{\mu\cdots\nu}\xi^\beta{}_{;\tau} \\ &+ A^{\alpha\cdots\beta}{}_{\tau\cdots\nu}\xi^\tau{}_{;\mu} + \cdots + A^{\alpha\cdots\beta}{}_{\mu\cdots\tau}\xi^\tau{}_{;\nu}.\end{aligned} \tag{9}$$

We now apply expression (9) to the case that A is a scalar σ, a four-vector V^μ and a tensor $A_{\mu\nu}$ respectively,

$$\hat{\sigma}(x) = \sigma(x) + \xi^\tau(x)\partial_\tau\sigma(x), \tag{10a}$$

$$\hat{V}^\mu = V^\mu + V^\mu{}_{;\tau}\xi^\tau - V^\tau\xi^\mu{}_{;\tau}, \tag{10b}$$

$$\hat{A}_{\mu\nu} = A_{\mu\nu} + A_{\mu\nu;\tau}\xi^\tau + A_{\tau\nu}\xi^\tau{}_{;\mu} + A_{\mu\tau}\xi^\tau{}_{;\nu}. \tag{10c}$$

For the metric tensor, $g_{\mu\nu}$ we find in particular, from expression (10c),

$$\hat{g}_{\mu\nu} = g_{\mu\nu} + \xi_{\mu;\nu} + \xi_{\nu;\mu}, \tag{11}$$

where we have used that the covariant derivative of the metric vanishes.

Our construction of gauge-invariant perturbations totally rest upon these expressions for hatted quantities. In case $\sigma(x)$ is some scalar quantity obeying the Einstein equations, $\sigma(x)$ can be divided in the usual way into a zeroth-order and a first-order part:

$$\sigma(x) \equiv \sigma_{(0)}(t) + \sigma_{(1)}(x), \tag{12}$$

where $\sigma_{(0)}(t)$ is some background quantity, and hence, not dependent on the spatial coordinates. Then (10a) becomes

$$\hat{\sigma}(x) = \sigma_{(0)}(t) + \sigma_{(1)}(x) + \xi^0(x)\partial_0\sigma_{(0)}(t) + \xi^\mu(x)\partial_\mu\sigma_{(1)}(x). \tag{13}$$

The last term, being a product of the first-order quantity $\xi^\mu(x)$ and the first-order quantity $\partial_\mu\sigma_{(1)}$, will be neglected. We thus find

$$\hat{\sigma}(x) = \sigma_{(0)}(t) + \hat{\sigma}_{(1)}(x), \tag{14}$$

with

$$\hat{\sigma}_{(1)}(x) \equiv \sigma_{(1)}(x) + \psi(x)\partial_0\sigma_{(0)}(t), \tag{15}$$

where we used (6). Thus, in gauge transformations of scalar quantities, only the zero component of the gauge functions need to be taken into account. Similarly, we find from (10b) and (11)

$$\hat{V}^\mu_{(1)}(x) = V^\mu_{(1)} + V^\mu_{(0);\tau}\xi^\tau - V^\tau_{(0)}\xi^\mu{}_{;\tau}, \tag{16}$$

and

$$\hat{g}_{(1)\mu\nu}(x) = g_{(1)\mu\nu}(x) + \xi_{\mu;\nu} + \xi_{\nu;\mu}. \tag{17}$$

The latter two expressions will be used later.

We are now in a position that we can conclude the proof of the statement that $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ are gauge-invariant. To that end, we now write down expression (15) once again, for another arbitrary scalar quantity $\omega(x)$ obeying the Einstein equations. We then find the analogue of expression (15)

$$\hat{\omega}_{(1)}(x) = \omega_{(1)}(x) + \psi(x)\partial_0\omega_{(0)}(t). \tag{18}$$

The left-hand sides of (15) and (18) give the value of the perturbation at the point with coordinates x with respect to the old coordinate system $\{x\}$; the right-hand sides of (15) and (18) contains quantities with the same values of the coordinates, x, but now with respect to the new coordinate system $\{\hat{x}\}$. Eliminating the function $\psi(x)$ from expressions (15) and (18) yields

$$\hat{\sigma}_{(1)}(x) - \frac{\partial_0\sigma_{(0)}(t)}{\partial_0\omega_{(0)}(t)}\hat{\omega}_{(1)}(x) = \sigma_{(1)}(x) - \frac{\partial_0\sigma_{(0)}(t)}{\partial_0\omega_{(0)}(t)}\omega_{(1)}(x). \tag{19}$$

In other words, the particular linear combination occurring in the right-hand side of (19) of any two scalar quantities ω and σ is gauge-invariant, and, hence, a possible candidate for a physical quantity.

The expressions (3) are precisely of the form (19). As a consequence, $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ are indeed invariant under the general infinitesimal coordinate transformation (4), i.e. they are gauge-invariant.

B. Unique Gauge-invariant Density Perturbations

Expression (19) is the key expression of this article as far as the scalar quantities $\varepsilon_{(1)}$ and $n_{(1)}$ are concerned. It tells us how to combine the scalar quantities occurring in the linearized Einstein equations in such a way that they become gauge independent. Expression (19) can be used to immediately derive the expressions (3) for the gauge-invariant energy and particle number densities.

In fact, let $U^{\mu}(x)$ be the four-velocity of the cosmological fluid. In Section IX C it is shown that in the linear theory of cosmological perturbations, defined by the background equations (154) and the perturbation equations (161) *only three independent scalars* play a role, namely

$$\varepsilon(x) = c^{-2}T^{\mu\nu}(x)U_{\mu}(x)U_{\nu}(x), \tag{20a}$$

$$n(x) = c^{-2}N^{\mu}(x)U_{\mu}(x), \tag{20b}$$

$$\theta(x) = c^{-1}U^{\mu}{}_{;\mu}(x), \tag{20c}$$

where

$$N^{\mu} \equiv nU^{\mu}, \tag{21}$$

is the cosmological particle current four-vector normalized according to $U^{\mu}U_{\mu} = c^2$. These scalars are divided according to

$$\varepsilon(x) = \varepsilon_{(0)}(t) + \varepsilon_{(1)}(x), \tag{22a}$$

$$n(x) = n_{(0)}(t) + n_{(1)}(x), \tag{22b}$$

$$\theta(x) = \theta_{(0)}(t) + \theta_{(1)}(x), \tag{22c}$$

where the background quantities $\varepsilon_{(0)}(t)$, $n_{(0)}(t)$ and $\theta_{(0)}(t)$ are solutions of the unperturbed Einstein equations. These quantities depend on the time coordinate t only. The relation (19) inspires us to consider the gauge-invariant combinations

$$\varepsilon^{\text{gi}}_{(1)}(x) \equiv \varepsilon_{(1)}(x) - \frac{\partial_0\varepsilon_{(0)}(t)}{\partial_0\omega_{(0)}(t)}\omega_{(1)}(x), \tag{23a}$$

$$n^{\text{gi}}_{(1)}(x) \equiv n_{(1)}(x) - \frac{\partial_0 n_{(0)}(t)}{\partial_0\omega_{(0)}(t)}\omega_{(1)}(x), \tag{23b}$$

$$\theta^{\text{gi}}_{(1)}(x) \equiv \theta_{(1)}(x) - \frac{\partial_0\theta_{(0)}(t)}{\partial_0\omega_{(0)}(t)}\omega_{(1)}(x). \tag{23c}$$

The question remains what to choose for ω in these three cases. In principle, we could choose for ω any of the following three scalar functions available in the theory, i.e. we could choose ε, n or θ. As we will show in Section XII, the only choice which satisfies the perturbed energy density constraint equation (161a) in the non-relativistic limit $v/c \to 0$ is

$$\omega = \theta. \tag{24}$$

This implies the expressions (3a) and (3b) for the energy and particle number density perturbations, as was to be shown. Using (24), we find from (23c)

$$\theta^{\text{gi}}_{(1)} \equiv \theta_{(1)} - \frac{\partial_0\theta_{(0)}}{\partial_0\theta_{(0)}}\theta_{(1)} = 0, \quad \theta_{(1)} \neq 0. \tag{25}$$

The physical interpretation of (25) is that, in first-order, the *global* expansion (20c) is not affected by a *local* perturbation in the energy density and particle number density. It should be emphasized here that (25) is *not* equivalent to the ‘uniform Hubble constant gauge’ of Bardeen [12], i.e. we do *not* impose the gauge condition $\theta_{(1)} = 0$. In contrast, $\theta^{\text{gi}}_{(1)} \equiv 0$ follows from the linearized Einstein equations, and, due to its gauge-invariance, *holds true in arbitrary systems of reference*. In other words, a ‘uniform Hubble function’ is inherent in a relativistic cosmological perturbation theory.

The fact that the expressions (3) for the gauge-invariant quantities $\varepsilon^{\text{gi}}_{(1)}$ and $n^{\text{gi}}_{(1)}$ are unique follows immediately from the background Einstein equations (154) and their perturbed counterparts (161). Consequently, these quantities cannot be chosen arbitrarily. In Section XII on the non-relativistic limit we show that $\varepsilon^{\text{gi}}_{(1)}$ is the perturbation to the energy density and $n^{\text{gi}}_{(1)}$ is the perturbation to the particle number density.

V. EINSTEIN EQUATIONS AND CONSERVATION LAWS IN SYNCHRONOUS COORDINATES

The system of evolution equations (197) or, equivalently (201), the main results of this article, are manifestly gauge-invariant. Therefore, one may use any convenient and suitable system of reference to derive these results.

The choice of a suitable coordinate system can be made as follows. It is well-known that in the Newtonian theory of gravity all possible space-time coordinate systems are *synchronous*, since time and space transformations (220) are decoupled in the Newtonian theory. Consequently, in order to show that our perturbation theory yields the Newtonian theory of gravity in the non-relativistic limit, it is obligatory to work in a synchronous system of reference. A second motivation to use synchronous coordinates is the fact that the background equations (154) are already given with respect to synchronous coordinates. Therefore, the evolution equations for scalar perturbations (161) turn out to be —in synchronous coordinates— simple extensions of the background equations. This fact has helped us to find the scalars (20) which play a key role in the construction of our manifestly gauge-invariant perturbation theory.

The name synchronous stems from the fact that surfaces with $t =$ constant are surfaces of simultaneity for observers at rest with respect to the synchronous coordinates, i.e. observers for which the three space coordinates x^i $(i = 1, 2, 3)$ remain constant. A synchronous system can be used for an arbitrary space-time manifold, not necessarily a homogeneous or homogeneous and isotropic one [23].

In a synchronous system of reference the line element for the metric has the form:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - g_{ij}(t,\boldsymbol{x})\mathrm{d}x^i\mathrm{d}x^j. \tag{26}$$

In a synchronous system, the coordinate t measures proper time along lines of constant x^i. From (26) we can read off that $(x^0 = ct)$:

$$g_{00}(t,\boldsymbol{x}) = 1, \quad g_{0i}(t,\boldsymbol{x}) = 0. \tag{27}$$

From the form of the line element in four-space (26) it follows that minus $g_{ij}(t,\boldsymbol{x})$, $(i = 1, 2, 3)$, is the metric of three-dimensional subspaces with constant t. Because of (27), knowing the three-geometry in all hyper-surfaces, is equivalent to knowing the geometry of space-time. The following abbreviations will prove useful when we rewrite the Einstein equations with respect to synchronous coordinates:

$$\varkappa_{ij} \equiv -\tfrac{1}{2}\dot{g}_{ij}, \quad \varkappa^i{}_j \equiv g^{ik}\varkappa_{kj}, \quad \varkappa^{ij} \equiv +\tfrac{1}{2}\dot{g}^{ij}, \tag{28}$$

where a dot denotes differentiation with respect to $x^0 = ct$. From (27)–(28) it follows that the connection coefficients of (four-dimensional) space-time

$$\Gamma^\lambda{}_{\mu\nu} = \tfrac{1}{2}g^{\lambda\rho}\left(g_{\rho\mu,\nu} + g_{\rho\nu,\mu} - g_{\mu\nu,\rho}\right), \tag{29}$$

are, in synchronous coordinates, given by

$$\Gamma^0{}_{00} = \Gamma^i{}_{00} = \Gamma^0{}_{i0} = \Gamma^0{}_{0i} = 0, \tag{30a}$$

$$\Gamma^0{}_{ij} = \varkappa_{ij}, \quad \Gamma^i{}_{0j} = \Gamma^i{}_{j0} = -\varkappa^i{}_j, \tag{30b}$$

$$\Gamma^k{}_{ij} = \tfrac{1}{2}g^{kl}\left(g_{li,j} + g_{lj,i} - g_{ij,l}\right). \tag{30c}$$

From (30c) it follows that the $\Gamma^k{}_{ij}$ are also the connection coefficients of (three-dimensional) subspaces of constant time.

The Ricci tensor $R_{\mu\nu} \equiv R^\lambda{}_{\mu\lambda\nu}$ is, in terms of the connection coefficients, given by

$$R_{\mu\nu} = \Gamma^\lambda{}_{\mu\nu,\lambda} - \Gamma^\lambda{}_{\mu\lambda,\nu} + \Gamma^\sigma{}_{\mu\nu}\Gamma^\lambda{}_{\lambda\sigma} - \Gamma^\sigma{}_{\mu\lambda}\Gamma^\lambda{}_{\nu\sigma}. \tag{31}$$

Upon substituting (30) into (31) one finds for the components of the Ricci tensor

$$R_{00} = \dot{\varkappa}^k{}_k - \varkappa^l{}_k\varkappa^k{}_l, \tag{32a}$$

$$R_{0i} = \varkappa^k{}_{k|i} - \varkappa^k{}_{i|k}, \tag{32b}$$

$$R_{ij} = \dot{\varkappa}_{ij} - \varkappa_{ij}\varkappa^k{}_k + 2\varkappa_{ik}\varkappa^k{}_j + {}^3\!R_{ij}, \tag{32c}$$

where the vertical bar in (32b) denotes covariant differentiation with respect to the metric g_{ij} of a three-dimensional subspace:

$$\varkappa^i{}_{j|k} \equiv \varkappa^i{}_{j,k} + \Gamma^i{}_{lk}\varkappa^l{}_j - \Gamma^l{}_{jk}\varkappa^i{}_l. \tag{33}$$

The quantities ${}^3R_{ij}$ in (32c) are found to be given by

$$ {}^3R_{ij} = \Gamma^k{}_{ij,k} - \Gamma^k{}_{ik,j} + \Gamma^l{}_{ij}\Gamma^k{}_{kl} - \Gamma^l{}_{ik}\Gamma^k{}_{jl}. \tag{34} $$

Hence, ${}^3R_{ij}$ is the Ricci tensor of the three-dimensional subspaces of constant time. For the components $R^\mu{}_\nu = g^{\mu\tau}R_{\tau\nu}$ of the Ricci tensor (32), we get

$$ R^0{}_0 = \dot{\varkappa}^k{}_k - \varkappa^l{}_k\varkappa^k{}_l, \tag{35a} $$
$$ R^0{}_i = \varkappa^k{}_{k|i} - \varkappa^k{}_{i|k}, \tag{35b} $$
$$ R^i{}_j = \dot{\varkappa}^i{}_j - \varkappa^i{}_j\varkappa^k{}_k + {}^3R^i{}_j, \tag{35c} $$

where we have used expressions (27)–(28).

The Einstein equations read

$$ G^{\mu\nu} - \Lambda g^{\mu\nu} = \kappa T^{\mu\nu}, \tag{36} $$

where $G^{\mu\nu}$, the Einstein tensor, is given by

$$ G^{\mu\nu} \equiv R^{\mu\nu} - \tfrac{1}{2}R^\alpha{}_\alpha g^{\mu\nu}. \tag{37} $$

In (36), Λ is a positive constant, the well-known cosmological constant. The constant κ is given by

$$ \kappa \equiv \frac{8\pi G}{c^4} = 2.0766 \times 10^{-43}\,\mathrm{m}^{-1}\,\mathrm{kg}^{-1}\,\mathrm{s}^2, \tag{38} $$

with $G = 6.6742 \times 10^{-11}\,\mathrm{m}^3\,\mathrm{kg}^{-1}\,\mathrm{s}^{-2}$ Newton's gravitational constant and $c = 2.99792458 \times 10^8\,\mathrm{m\,s}^{-1}$ the speed of light. In view of the Bianchi identities one has $G^{\mu\nu}{}_{;\nu} = 0$, hence, since $g^{\mu\nu}{}_{;\nu} = 0$, the source term $T^{\mu\nu}$ of the Einstein equations must fulfill the properties

$$ T^{\mu\nu}{}_{;\nu} = 0. \tag{39} $$

These equations are the energy-momentum conservation laws.

In order to derive simultaneously the background and first-order equations, we rewrite the Einstein equations (36) in an alternative form, using mixed upper and lower indices:

$$ R^\mu{}_\nu = \kappa(T^\mu{}_\nu - \tfrac{1}{2}\delta^\mu{}_\nu T^\alpha{}_\alpha) - \Lambda\delta^\mu{}_\nu. \tag{40} $$

Upon substituting the components (35) into the Einstein equations (40), and eliminating the time derivative of $\varkappa^k{}_k$ from the $R^0{}_0$-equation with the help of the $R^i{}_j$-equations, the Einstein equations can be cast in the form

$$ (\varkappa^k{}_k)^2 - {}^3R - \varkappa^k{}_l\varkappa^l{}_k = 2(\kappa T^0{}_0 + \Lambda), \tag{41a} $$
$$ \varkappa^k{}_{k|i} - \varkappa^k{}_{i|k} = \kappa T^0{}_i, \tag{41b} $$
$$ \dot{\varkappa}^i{}_j - \varkappa^i{}_j\varkappa^k{}_k + {}^3R^i{}_j = \kappa(T^i{}_j - \tfrac{1}{2}\delta^i{}_j T^\mu{}_\mu) - \Lambda\delta^i{}_j, \tag{41c} $$

where

$$ {}^3R \equiv g^{ij}\,{}^3R_{ij} = {}^3R^k{}_k, \tag{42} $$

is the curvature scalar of the three-dimensional subspaces of constant time. The (differential) equations (41c) are the so-called dynamical Einstein equations: they define the evolution (of the time derivative) of the (spatial part of the) metric. The (algebraic) equations (41a) and (41b) are constraint equations: they relate the initial conditions, and, once these are satisfied at one time, they are satisfied automatically at all times.

The right-hand side of equations (41) contain the components of the energy momentum tensor $T_{\mu\nu}$, which, for a perfect fluid, are given by

$$ T^\mu{}_\nu = (\varepsilon + p)u^\mu u_\nu - p\delta^\mu{}_\nu, \tag{43} $$

where $u^\mu(t,\boldsymbol{x}) = c^{-1}U^\mu(t,\boldsymbol{x})$ is the hydrodynamic fluid four-velocity normalized to unity ($u^\mu u_\mu = 1$), $\varepsilon(t,\boldsymbol{x})$ the energy density and $p(t,\boldsymbol{x})$ the pressure at a point $(t,\boldsymbol{x})$ in space-time. In this expression we neglect terms containing the shear and volume viscosity, and other terms related to irreversible processes. The equation of state for the pressure

$$ p = p(n, \varepsilon), \tag{44} $$

where $n(t, \boldsymbol{x})$ is the particle number density at a point $(t, \boldsymbol{x})$ in space-time, is supposed to be a given function of n and ε (see also Appendix A for equations of state in alternative forms).

As stated above already, the Einstein equations (41a) and (41b) are constraint equations to the Einstein equations (41c) only: they tell us what relations should exist between the initial values of the various unknown functions, in order that the Einstein equations be solvable. In the following, we shall suppose that these conditions are satisfied. Thus we are left with the nine equations (41c), of which, because of the symmetry of g_{ij}, only six are independent. These six equations, together with the four equations (39) constitute a set of ten equations for the eleven $(6+3+1+1)$ independent quantities g_{ij}, u^i, ε and n. The eleventh equation needed to close the system of equations is the particle number conservation law $N^{\mu}{}_{;\mu} = 0$. Using (21), we get

$$(nu^{\mu})_{;\mu} = 0, \tag{45}$$

where a semicolon denotes covariant differentiation with respect to the metric tensor $g_{\mu\nu}$. This equation can be rewritten in terms of the fluid expansion scalar defined by expression (20c). Using (30), we can rewrite the four-divergence (20c) in the form

$$\theta = \dot{u}^0 - \varkappa^k{}_k u^0 + \vartheta, \tag{46}$$

where the three-divergence ϑ is given by

$$\vartheta \equiv u^k{}_{|k}. \tag{47}$$

Using now expressions (20c), (30), (33) and (46), the four energy-momentum conservation laws (39) and the particle number conservation law (45) can be rewritten as

$$\dot{T}^{00} + T^{0k}{}_{|k} + \varkappa^k{}_l T^l{}_k - \varkappa^k{}_k T^{00} = 0, \tag{48a}$$
$$\dot{T}^{i0} + T^{ik}{}_{|k} - 2\varkappa^i{}_k T^{k0} - \varkappa^k{}_k T^{i0} = 0, \tag{48b}$$

and

$$\dot{n}u^0 + n_{,k}u^k + n\theta = 0, \tag{49}$$

respectively. Since T^{0i} is a vector and T^{ij} is a tensor with respect to coordinate transformations in a subspace of constant time, and, hence, are tensorial quantities in this three-dimensional subspace, we could use in (48) a bar to denote covariant differentiation with respect to the metric $g_{ij}(t, \boldsymbol{x})$ of such a subspace of constant time t.

The Einstein equations (41) and conservation laws (48) and (49) describe a universe filled with a perfect fluid and with a positive cosmological constant. The fluid pressure p is described by an equation of state of the form (44): in this stage we only need that it is some function of the particle number density n and the energy density ε.

We have now rewritten the Einstein equations and conservation laws in such a way that one can easily derive the background and perturbation equations.

VI. ZEROTH- AND FIRST-ORDER EQUATIONS FOR THE FLRW UNIVERSE

We will now limit the discussion to a particular class of universes, namely the collection of universes that, apart from a small, local perturbation in space-time, are homogeneous and isotropic, the Friedmann-Lemaître-Robertson-Walker (FLRW) universes.

We expand all quantities Q in the form of series, and derive, recursively, equations for the successive terms of these series. We will distinguish the successive terms of a series by a sub-index between brackets:

$$Q = Q_{(0)} + \eta Q_{(1)} + \eta^2 Q_{(2)} + \cdots, \tag{50}$$

where the sub-index zero refers to quantities of the unperturbed, homogeneous and isotropic FLRW universe.

In expression (50) η ($\eta \equiv 1$) is a bookkeeping parameter, the function of which is to enable us in actual calculations to easily distinguish between the terms of different orders.

A. Zeroth-order Quantities

This section is concerned with the background or zeroth-order quantities occurring in the Einstein equations. All results of this section are standard [35], and given here only to fix the notation unambiguously.

For a FLRW universe, the background metric $g_{(0)\mu\nu}$ is given by

$$g_{(0)00}(t,\boldsymbol{x}) = 1, \quad g_{(0)0i}(t,\boldsymbol{x}) = 0, \tag{51a}$$

$$g_{(0)ij}(t,\boldsymbol{x}) = -a^2(t)\tilde{g}_{ij}(\boldsymbol{x}), \quad g^{ij}_{(0)}(t,\boldsymbol{x}) = -\frac{1}{a^2(t)}\tilde{g}^{ij}(\boldsymbol{x}). \tag{51b}$$

where $\tilde{g}_{ij}(\boldsymbol{x})$ is the metric of a three-dimensional maximally symmetric subspace:

$$\tilde{g}_{ij} = \mathrm{diag}\left(\frac{1}{1-kr^2}, r^2, r^2\sin^2\varpi\right), \quad k = 0, \pm 1. \tag{52}$$

The minus sign in (51b) has been introduced in order to switch from the conventional four-dimensional space-time with signature $(+,-,-,-)$ to the conventional three-dimensional spatial metric with signature $(+,+,+)$.

All background scalars depend on time only. Furthermore, four-vectors have vanishing spatial components. Since u^μ is a unit vector we have

$$u^\mu_{(0)} = \delta^\mu{}_0. \tag{53}$$

The time derivative of the three-part of the metric $g_{(0)ij}$, $\varkappa_{(0)ij}$, may be expressed in the Hubble function $\mathcal{H}(t) \equiv (\mathrm{d}a/\mathrm{d}t)/a(t)$. We prefer to use a function

$$H(t) = \frac{\mathcal{H}(t)}{c}, \tag{54}$$

which we will call Hubble function also. Recalling that a dot denotes differentiation with respect to ct, we have

$$H \equiv \frac{\dot{a}}{a}. \tag{55}$$

Substituting the expansion (50) into the definitions (28), we obtain

$$\varkappa_{(0)ij} = -Hg_{(0)ij}, \;\; \varkappa^i_{(0)j} = -H\delta^i{}_j, \;\; \varkappa^{ij}_{(0)} = -Hg^{ij}_{(0)}, \tag{56}$$

where we considered only terms up to the zeroth-order in the bookkeeping parameter η.

Similarly, with (46), (47), (50), (53) and (56) we find for the background fluid expansion scalar, $\theta_{(0)}$, and the three-divergence, $\vartheta_{(0)}$,

$$\theta_{(0)} = 3H, \quad \vartheta_{(0)} = 0. \tag{57}$$

These quantities and their first-order counterparts will play and important role in our perturbation theory.

Using (43), (50), (51) and (53) we find for the components of the energy momentum tensor

$$T^0_{(0)0} = \varepsilon_{(0)}, \quad T^i_{(0)0} = 0, \quad T^i_{(0)j} = -p_{(0)}\delta^i{}_j, \tag{58}$$

where the background pressure $p_{(0)}$ is given by the equation of state (44), which, for the background pressure, is defined by

$$p_{(0)} = p_{(0)}(n_{(0)}, \varepsilon_{(0)}). \tag{59}$$

The background three-dimensional Ricci tensor, (34), is given by

$${}^3R_{(0)ij} = \Gamma^k_{(0)ij,k} - \Gamma^k_{(0)ik,j} + \Gamma^l_{(0)ij}\Gamma^k_{(0)kl} - \Gamma^l_{(0)ik}\Gamma^k_{(0)jl}, \tag{60}$$

where the connection coefficients $\Gamma^k_{(0)ij}$ are given by

$$\Gamma^k_{(0)ij} = \tfrac{1}{2}g^{kl}_{(0)}\left(g_{(0)li,j} + g_{(0)lj,i} - g_{(0)ij,l}\right), \tag{61}$$

where $g^{ij}_{(0)}$ and $g_{(0)ij}$ depend on time. Hence, the connection coefficients $\Gamma^k_{(0)ij}$ are equal to the connection coefficients $\tilde{\Gamma}^k{}_{ij}$ of the metric $\tilde{g}_{ij}$:

$$\Gamma^k_{(0)ij} = \tilde{\Gamma}^k{}_{ij} \equiv \tfrac{1}{2}\tilde{g}^{kl}\left(\tilde{g}_{li,j} + \tilde{g}_{lj,i} - \tilde{g}_{ij,l}\right). \tag{62}$$

Therefore, they do not depend on time.

Substituting (51) and (52) into (60), combined with (62), we find

$$^3R_{(0)ij} = 2k\tilde{g}_{ij}. \tag{63}$$

From (63) we have

$$^3R^i_{(0)j}(t) = -\frac{2k}{a^2(t)}\delta^i{}_j, \tag{64}$$

implying that the zeroth-order curvature scalar $^3R_{(0)} = g^{ij}_{(0)}\,{}^3R_{(0)ij}$ is given by

$$^3R_{(0)}(t) = -\frac{6k}{a^2(t)}. \tag{65}$$

This quantity and its perturbed counterpart will play an important role in our perturbation theory. Note, that in view of our choice of the metric $(+,-,-,-)$, spaces of positive curvature k have a negative curvature scalar $^3R_{(0)}$.

Thus, we have found all background quantities.

B. First-order Quantities

In this section we express all quantities occurring in the Einstein equations in terms of zeroth- and first-order quantities. The equations of state for the energy and pressure, $\varepsilon(n,T)$ and $p(n,T)$, are not specified yet.

Upon substituting the series (50) into the normalization condition $u^\mu u_\mu = 1$, one finds, equating equal powers of the bookkeeping parameter η,

$$u^0_{(1)} = 0, \tag{66}$$

for the first-order perturbation to the four-velocity. Writing the inverse of $g_{kl} = g_{(0)kl} + \eta g_{(1)kl} + \cdots$ as

$$g^{kl} = g^{kl}_{(0)} + \eta g^{kl}_{(1)} + \cdots, \tag{67}$$

where $g^{kl}_{(0)}$ is the inverse, of $g_{(0)kl}$, (51b), we find

$$g^{kl}_{(1)} = -g^{ki}_{(0)}g^{lj}_{(0)}g_{(1)ij}, \tag{68}$$

and

$$g^k_{(1)i} = -g^{kl}_{(0)}g_{(1)li}. \tag{69}$$

It is convenient to introduce

$$h_{ij} \equiv -g_{(1)ij}, \tag{70}$$

so that

$$h^{ij} = g^{ij}_{(1)}, \quad h^i{}_j = g^{ik}_{(0)}h_{kj}. \tag{71}$$

For the time derivative of the first-order perturbations to the metric, $\varkappa_{(1)ij}$ (28), we get

$$\varkappa_{(1)ij} = \tfrac{1}{2}\dot{h}_{ij}, \quad \varkappa^i_{(1)j} = \tfrac{1}{2}\dot{h}^i{}_j, \quad \varkappa^{ij}_{(1)} = \tfrac{1}{2}\dot{h}^{ij}. \tag{72}$$

The first-order perturbation $\theta_{(1)}$ to the fluid expansion scalar θ (46), can be found in the same way. Using (50) and (53) one arrives at

$$\theta_{(1)} = \vartheta_{(1)} - \tfrac{1}{2}\dot{h}^k{}_k, \tag{73}$$

where we used (66) and (72). This expression will play an important role in the derivation of the first-order perturbation equations in Section VIII. The first-order perturbation $\vartheta_{(1)}$ to the three-divergence ϑ (47), is

$$\vartheta_{(1)} = u^k_{(1)|k}, \tag{74}$$

where we have used that

$$(u^k{}_{|k})_{(1)} = u^k_{(1)|k}, \tag{75}$$

which is a consequence of $\Gamma^k_{(1)lk} u^l_{(0)} = 0$ as follows from (53).

Upon substituting the series expansion (50), and into (43) and equating equal powers of η, one finds for the first-order perturbation to the energy-momentum tensor

$$T^0_{(1)0} = \varepsilon_{(1)}, \tag{76a}$$
$$T^i_{(1)0} = (\varepsilon_{(0)} + p_{(0)}) u^i_{(1)}, \tag{76b}$$
$$T^i_{(1)j} = -p_{(1)} \delta^i{}_j, \tag{76c}$$

where we have used (53) and (66). The first-order perturbation to the pressure is related to $\varepsilon_{(1)}$ and $n_{(1)}$ by the first-order perturbation to the equation of state (44). We have

$$p_{(1)} = p_n n_{(1)} + p_\varepsilon \varepsilon_{(1)}, \tag{77}$$

where p_n and p_ε are the partial derivatives of $p(n, \varepsilon)$ with respect to n and ε,

$$p_n \equiv \left(\frac{\partial p}{\partial n}\right)_\varepsilon, \quad p_\varepsilon \equiv \left(\frac{\partial p}{\partial \varepsilon}\right)_n. \tag{78}$$

Since we consider only first-order quantities, the partial derivatives are functions of the background quantities only, i.e.

$$p_n = p_n(n_{(0)}, \varepsilon_{(0)}), \quad p_\varepsilon = p_\varepsilon(n_{(0)}, \varepsilon_{(0)}). \tag{79}$$

Using (30c) and the series (50), we find for the first-order perturbations of the connection coefficients

$$\Gamma^k_{(1)ij} = -g^{kl}_{(0)} g_{(1)lm} \Gamma^m_{(0)ij} + \tfrac{1}{2} g^{kl}_{(0)} \left(g_{(1)li,j} + g_{(1)lj,i} - g_{(1)ij,l}\right). \tag{80}$$

The first-order perturbations $\Gamma^k_{(1)ij}$ (80), occurring in the non-tensor $\Gamma^k{}_{ij}$, happen to be expressible as a tensor. Indeed, using (70), one can rewrite (80) in the form

$$\Gamma^k_{(1)ij} = -\tfrac{1}{2} g^{kl}_{(0)} (h_{li|j} + h_{lj|i} - h_{ij|l}). \tag{81}$$

Using the expansion (50) for ${}^3R_{ij}$ and $\Gamma^k{}_{ij}$, one finds for the first-order perturbation to the Ricci tensor (34)

$${}^3R_{(1)ij} = \Gamma^k_{(1)ij,k} - \Gamma^k_{(1)ik,j} + \Gamma^l_{(0)ij}\Gamma^k_{(1)kl} + \Gamma^l_{(1)ij}\Gamma^k_{(0)kl} - \Gamma^l_{(0)ik}\Gamma^k_{(1)jl} - \Gamma^l_{(1)ik}\Gamma^k_{(0)jl}, \tag{82}$$

which can be rewritten in the compact form

$${}^3R_{(1)ij} = \Gamma^k_{(1)ij|k} - \Gamma^k_{(1)ik|j}. \tag{83}$$

By substituting (81) into (83), one can express the first-order perturbation to the Ricci tensor of the three-dimensional subspace in terms of the perturbation to the metric and its covariant derivatives:

$${}^3R_{(1)ij} = -\tfrac{1}{2} g^{kl}_{(0)} (h_{li|j|k} + h_{lj|i|k} - h_{ij|l|k} - h_{lk|i|j}). \tag{84}$$

The perturbation ${}^3R^i_{(1)j}$ is given by

$${}^3R^i_{(1)j} \equiv (g^{ip}\,{}^3R_{pj})_{(1)} = g^{ip}_{(0)}\,{}^3R_{(1)pj} + \tfrac{1}{3}\,{}^3R_{(0)} h^i{}_j, \tag{85}$$

where we have used (51), (63), (65) and (71). Upon substituting (84) into (85) we get

$${}^3R^i_{(1)j} = -\tfrac{1}{2} g^{ip}_{(0)} (h^k{}_{p|j|k} + h^k{}_{j|p|k} - h^k{}_{k|p|j}) + \tfrac{1}{2} g^{kl}_{(0)} h^i{}_{j|k|l} + \tfrac{1}{3}\,{}^3R_{(0)} h^i{}_j. \tag{86}$$

Taking $i = j$ in (86) and summing over the repeated index, we find for the first-order perturbation to the curvature scalar of the three-dimensional spaces

$$
{}^3R_{(1)} = g^{ij}_{(0)}\left(h^k{}_{k|i|j} - h^k{}_{i|j|k}\right) + \tfrac{1}{3}\,{}^3R_{(0)} h^k{}_k. \tag{87}
$$

This expression will play an important role in the evolution of density perturbations, equations (195), and in the non-relativistic limit in Section XII.

We thus have expressed all quantities occurring in the relevant dynamical equations, i.e. the system of equations formed by the Einstein equations combined with the conservation laws, in terms of zeroth- and first-order quantities to be solved from these equations. In the Sections VI C and VI D below we derive the background and first-order evolution equations respectively. To that end we substitute the series (50) into the Einstein equations (41) and conservation laws (48) and (49). By equating the powers of η^0, η^1, ..., we obtain the zeroth-order, the first-order and higher order dynamical equations, constraint equations and conservation laws. We will carry out this scheme for the zeroth- and first-order equations only.

C. Zeroth-order Equations

With the help of Section VI A and the series (50) we now can find from the Einstein equations (41) and conservation laws (48) and (49) the zeroth-order Einstein equations and the conservation laws. Furthermore, in view of the symmetry induced by the isotropy, it is possible to switch from the six quantities g_{ij} and the six quantities $\varkappa_{ij}$ to the curvature ${}^3R_{(0)}(t)$ and the Hubble function $H(t)$ only.

1. *Einstein Equations*

Upon substituting (56) and (58) into the $(0,0)$-component of the Einstein equations, (41a), one finds

$$
3H^2 - \tfrac{1}{2}\,{}^3R_{(0)} = \kappa\varepsilon_{(0)} + \Lambda. \tag{88}
$$

The $(0,i)$-components of the Einstein equations, (41b), are identically fulfilled, as follows from (56) and (58). We thus are left with the six (i,j)-components of the Einstein equations, (41c). In view of (56), (58) and (64) we find that $\varkappa^1_{(0)1} = \varkappa^2_{(0)2} = \varkappa^3_{(0)3}$, $T^1_{(0)1} = T^2_{(0)2} = T^3_{(0)3}$ and ${}^3R^1_{(0)1} = {}^3R^2_{(0)2} = {}^3R^3_{(0)3}$, whereas for $i \neq j$ these quantities vanish. Hence, the six (i,j)-components reduce to one equation,

$$
\dot{H} = -3H^2 + \tfrac{1}{3}\,{}^3R_{(0)} + \tfrac{1}{2}\kappa(\varepsilon_{(0)} - p_{(0)}) + \Lambda. \tag{89}
$$

In equations (88) and (89) the background curvature ${}^3R_{(0)}$ is given by (65). It is, however, of convenience to determine this quantity from a differential equation. Eliminating $a(t)$ from expressions (55) and (65) we obtain

$$
{}^3\dot{R}_{(0)} + 2H\,{}^3R_{(0)} = 0, \tag{90}
$$

where the initial value ${}^3R_{(0)}(t_0)$ is given by

$$
{}^3R_{(0)}(t_0) = -\frac{6k}{a^2(t_0)}, \tag{91}
$$

in accordance with (65). It should be emphasized that equation (90) is not an Einstein equation, since it is equivalent to expression (65). It will be used here as an ancillary relation.

2. *Conservation Laws*

Upon substituting (56) and (58) into the 0-component of the conservation law, (48a), one finds

$$
\dot{\varepsilon}_{(0)} + 3H(\varepsilon_{(0)} + p_{(0)}) = 0, \tag{92}
$$

which is the relativistic background continuity equation. The background momentum conservation laws (i.e. the background relativistic Euler equations) are identically satisfied, as follows by substituting (56) and (58) into the spatial components of the conservation laws, (48b).

The background particle number density conservation law can be found by substituting (53) and (57) into equation (49). One gets

$$\dot{n}_{(0)} + 3Hn_{(0)} = 0. \tag{93}$$

It can be shown by differentiating the constraint equation (88) with respect to time that the general solution of the system (88) and (90)–(93) is also a solution of the dynamical equation (89). Consequently, equation (89) need not be considered anymore. This concludes the derivation of the background equations.

D. First-order Equations

In this section we derive the first-order perturbation equations from the Einstein equations (41) and conservation laws (48) and (49). The procedure is, by now, completely standard. We use the series expansion (50) in η for the various quantities occurring in the Einstein equations and conservation laws of energy-momentum, and we equate the coefficients linear in η to obtain the 'linearized' or first-order equations.

1. *Einstein Equations*

Using the series expansions (50) for ${}^3R$, $\varkappa^i{}_j$ and $T^0{}_0$, in the $(0,0)$-component of the constraint equation (41a), one finds

$$2\varkappa^k_{(0)k}\varkappa^l_{(1)l} - {}^3R_{(1)} - 2\varkappa^k_{(0)l}\varkappa^l_{(1)k} = 2\kappa T^0_{(1)0}. \tag{94}$$

With the zeroth-order expressions (56), the abbreviations (72) and the expression for $T^0_{(1)0}$, (76), we may rewrite this equation in the form

$$H\dot{h}^k{}_k + \tfrac{1}{2}\,{}^3R_{(1)} = -\kappa\varepsilon_{(1)}. \tag{95}$$

Using the series expansion (50) for $\varkappa^i{}_j$ and $T^0{}_i$, we find for the $(0,i)$-components of the constraint equations (41b)

$$\varkappa^k_{(1)k|i} - \varkappa^k_{(1)i|k} = \kappa T^0_{(1)i}, \tag{96}$$

where we noted that

$$(\varkappa^i{}_{j|k})_{(1)} = \varkappa^i_{(1)j|k}, \tag{97}$$

which is a consequence of $\Gamma^i_{(1)lk}\varkappa^l_{(0)j} - \Gamma^l_{(1)jk}\varkappa^i_{(0)l} = 0$, which, in turn, is a direct consequence of (56). From (72) and (76) we find

$$\dot{h}^k{}_{k|i} - \dot{h}^k{}_{i|k} = 2\kappa(\varepsilon_{(0)} + p_{(0)})u_{(1)i}. \tag{98}$$

Finally, we consider the (i,j)-components of the Einstein equations (41c). Using the series expansion (50) for $\varkappa^i{}_j$, $T^i{}_j$ and ${}^3R^i{}_j$, we find

$$\dot{\varkappa}^i_{(1)j} - \varkappa^i_{(1)j}\varkappa^k_{(0)k} - \varkappa^i_{(0)j}\varkappa^k_{(1)k} + {}^3R^i_{(1)j} = \kappa(T^i_{(1)j} - \tfrac{1}{2}\delta^i{}_j T^\mu_{(1)\mu}). \tag{99}$$

With (56), (72) and (76), we get

$$\ddot{h}^i{}_j + 3H\dot{h}^i{}_j + \delta^i{}_j H\dot{h}^k{}_k + 2\,{}^3R^i_{(1)j} = -\kappa\delta^i{}_j(\varepsilon_{(1)} - p_{(1)}), \tag{100}$$

where ${}^3R^i_{(1)j}$ is given by expression (86).

Note that the first-order equations (95) and (100) are independent of the cosmological constant Λ: the effect of the non-zero cosmological constant is accounted for by the zeroth-order quantities [cf. equations (88) and (89)].

2. Conservation Laws

We now consider the energy conservation law (48a). With the help of the series expansion (50) for $\varkappa^i{}_j$ and $T^\mu{}_\nu$, one finds for the first-order equation

$$\dot{T}^{00}_{(1)} + T^{0k}_{(1)}{}_{|k} + \varkappa^k_{(0)l} T^l_{(1)k} + \varkappa^k_{(1)l} T^l_{(0)k} - \varkappa^k_{(0)k} T^{00}_{(1)} - \varkappa^k_{(1)k} T^{00}_{(0)} = 0, \tag{101}$$

where we have used that for a three-vector T^{0k} we have

$$(T^{0k}{}_{|k})_{(1)} = T^{0k}_{(1)}{}_{|k}, \tag{102}$$

see expression (75). Employing (56), (58), (72)–(74) and (76) we arrive at the first-order energy conservation law

$$\dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + (\varepsilon_{(0)} + p_{(0)})\theta_{(1)} = 0. \tag{103}$$

Next, we consider the momentum conservation laws (48b). With the series expansion (50) for $\varkappa^i{}_j$ and $T^{\mu\nu}$, we find for the first-order momentum conservation law

$$\dot{T}^{i0}_{(1)} + (T^{ik}{}_{|k})_{(1)} - 2\varkappa^i_{(0)k} T^{k0}_{(1)} - 2\varkappa^i_{(1)k} T^{k0}_{(0)} - \varkappa^k_{(0)k} T^{i0}_{(1)} - \varkappa^k_{(1)k} T^{i0}_{(0)} = 0. \tag{104}$$

Using that

$$(T^{ik}{}_{|k})_{(1)} = -g^{ik}_{(0)} p_{(1)|k}, \tag{105}$$

and expressions (56), (58), (72) and (76) we arrive at

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\Big[(\varepsilon_{(0)} + p_{(0)}) u^i_{(1)}\Big] - g^{ik}_{(0)} p_{(1)|k} + 5H(\varepsilon_{(0)} + p_{(0)}) u^i_{(1)} = 0, \tag{106}$$

where we have also used that the covariant derivative of $g^{ij}_{(0)}$ vanishes: $g^{ij}_{(0)|k} = 0$.

Finally, we consider the particle number density conservation law (49). With the series expansion (50) for n, θ, and u^μ, it follows that the first-order equation reads

$$\dot{n}_{(0)} u^0_{(1)} + \dot{n}_{(1)} u^0_{(0)} + n_{(0),k} u^k_{(1)} + n_{(1),k} u^k_{(0)} + n_{(0)}\theta_{(1)} + n_{(1)}\theta_{(0)} = 0. \tag{107}$$

With the help of (53), (57) and (66) we find for the first-order particle number conservation law

$$\dot{n}_{(1)} + 3H n_{(1)} + n_{(0)}\theta_{(1)} = 0. \tag{108}$$

This concludes the derivation of the basic perturbation equations.

3. Summary

In the preceding two Sections VI D 1 and VI D 2 we have found the equations which, basically, describe the perturbations in a FLRW universe, in first-order approximation. They are equations (95), (98), (100), (103), (106) and (108). For convenience we repeat them here

$$
\begin{aligned}
\text{Constraints:}\quad & H\dot{h}^k{}_k + \tfrac{1}{2}\,{}^3\!R_{(1)} = -\kappa\varepsilon_{(1)}, && (109\text{a})\\
& \dot{h}^k{}_{k|i} - \dot{h}^k{}_{i|k} = 2\kappa(\varepsilon_{(0)} + p_{(0)}) u_{(1)i}, && (109\text{b})\\
\text{Evolution:}\quad & \ddot{h}^i{}_j + 3H\dot{h}^i{}_j + \delta^i{}_j H\dot{h}^k{}_k + 2\,{}^3\!R^i_{(1)j} = -\kappa\delta^i{}_j(\varepsilon_{(1)} - p_{(1)}), && (109\text{c})\\
\text{Conservation:}\quad & \dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + (\varepsilon_{(0)} + p_{(0)})\theta_{(1)} = 0, && (109\text{d})\\
& \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\Big[(\varepsilon_{(0)} + p_{(0)}) u^i_{(1)}\Big] - g^{ik}_{(0)} p_{(1)|k} + 5H(\varepsilon_{(0)} + p_{(0)}) u^i_{(1)} = 0, && (109\text{e})\\
& \dot{n}_{(1)} + 3H n_{(1)} + n_{(0)}\theta_{(1)} = 0, && (109\text{f})
\end{aligned}
$$

where $\theta_{(1)}$, ${}^3\!R^i_{(1)j}$ and ${}^3\!R_{(1)}$ are given by (73), (86) and (87) respectively. Hence, the equations (109) essentially are fifteen equations for the eleven $(6+3+1+1)$ unknown functions $h^i{}_j$, $u^i_{(1)}$, $\varepsilon_{(1)}$ and $n_{(1)}$. The pressure $p_{(0)}$ is given by an equation of state (59), and the perturbation to the pressure, $p_{(1)}$, is given by (77). The system of equations is not over-determined, however, since the four equations (109a) and (109b) are only conditions on the initial values. These initial value conditions are fulfilled for all times t automatically if they are satisfied at some (initial) time $t = t_0$.

VII. CLASSIFICATION OF THE SOLUTIONS OF FIRST-ORDER

It is well-known that, for *flat* FLRW universes, the set of linear equations (109) can be divided into three *independent* sets of equations, which, together, are equivalent to the original set. In this section we show that this can also be done for the open and closed FLRW universes. We will refer to these sets by their usual names of scalar, vector and tensor perturbation equations. We will show that the vector and tensor perturbations do not, in first-order, contribute to the physical perturbations $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$. As a consequence, we only need, for our problem, the set of equations which are related to the scalar perturbations. By considering only the scalar part of the full set of perturbation equations we are able to cast the perturbation equations into a set which is directly related to the physical perturbations $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$. This is the subject of Section VIII.

At the basis of the replacement of one set (109) by three sets of equations stands a theorem proved by York [36] and Stewart [37], which states that a symmetric second rank tensor can be divided into three irreducible pieces, and that a vector can be divided into two irreducible pieces. Here, we will use this general theorem to obtain equations for the scalar irreducible parts of the tensors $h^i{}_j$ and ${}^3R^i_{(1)j}$ and the vector $\boldsymbol{u}_{(1)}$, namely $h^i_{\parallel j}$, ${}^3R^i_{(1)\parallel j}$ and $\boldsymbol{u}_{(1)\parallel}$.

For the perturbation to the metric, a symmetric second rank tensor we have in particular

$$h^i{}_j = h^i_{\parallel j} + h^i_{\perp j} + h^i_{*j}, \tag{110}$$

where, according to the theorem of Stewart [37], the irreducible constituents $h^i_{\parallel j}$, $h^i_{\perp j}$ and h^i_{*j} have the properties

$$h^i_{\parallel j} = \frac{2}{c^2}(\phi\delta^i{}_j + \zeta^{|i}{}_{|j}), \tag{111a}$$

$$h^k_{\perp k} = 0, \tag{111b}$$

$$h^k_{*k} = 0, \quad h^k_{*i|k} = 0, \tag{111c}$$

with $\phi(t,\boldsymbol{x})$ and $\zeta(t,\boldsymbol{x})$ arbitrary functions. The contravariant derivative $A^{|i}$ is defined as $g^{ij}_{(0)}A_{|j}$. The functions $h^i_{\parallel j}$, $h^i_{\perp j}$ and h^i_{*j} correspond to scalar, vector and tensor perturbations respectively.

In the same way, the perturbation to the Ricci tensor can be decomposed into irreducible components, i.e.

$${}^3R^i_{(1)j} = {}^3R^i_{(1)\parallel j} + {}^3R^i_{(1)\perp j} + {}^3R^i_{(1)*j}. \tag{112}$$

The tensors ${}^3R^i_{(1)\parallel j}$, ${}^3R^i_{(1)\perp j}$ and ${}^3R^i_{(1)*j}$ have the properties comparable to (111), i.e.

$${}^3R^i_{(1)\parallel j} = \frac{2}{c^2}(\gamma\delta^i{}_j + \pi^{|i}{}_{|j}), \tag{113a}$$

$${}^3R^k_{(1)\perp k} = 0, \tag{113b}$$

$${}^3R^k_{(1)*k} = 0, \quad {}^3R^k_{(1)*i|k} = 0, \tag{113c}$$

where $\gamma(t,\boldsymbol{x})$ and $\pi(t,\boldsymbol{x})$ are two arbitrary functions. By now using (86) for each of the irreducible parts we find

$$\begin{aligned}{}^3R^i_{(1)\parallel j} =& \frac{1}{c^2}\Big[\phi^{|i}{}_{|j} + \delta^i{}_j\phi^{|k}{}_{|k} - \zeta^{|k|i}{}_{|j|k} - \zeta^{|k}{}_{|j}{}^{|i}{}_{|k} \\ &+ \zeta^{|k}{}_{|k}{}^{|i}{}_{|j} + \zeta^{|i}{}_{|j}{}^{|k}{}_{|k} + \tfrac{2}{3}\,{}^3R_{(0)}(\delta^i{}_j\phi + \zeta^{|i}{}_{|j})\Big],\end{aligned} \tag{114a}$$

$${}^3R^i_{(1)\perp j} = -\tfrac{1}{2}g^{ip}_{(0)}(h^k_{\perp p|j|k} + h^k_{\perp j|p|k}) + \tfrac{1}{2}g^{kl}_{(0)}h^i_{\perp j|k|l} + \tfrac{1}{3}\,{}^3R_{(0)}h^i_{\perp j}, \tag{114b}$$

$${}^3R^i_{(1)*j} = -\tfrac{1}{2}g^{ip}_{(0)}(h^k_{*p|j|k} + h^k_{*j|p|k}) + \tfrac{1}{2}g^{kl}_{(0)}h^i_{*j|k|l} + \tfrac{1}{3}\,{}^3R_{(0)}h^i_{*j}. \tag{114c}$$

Combining expressions (113b) and (114b) it follows that $h^i_{\perp j}$ has the property

$$h^{kl}_{\perp|k|l} = 0, \tag{115}$$

in addition to the property (111b). In Section VII B we show that this additional condition is needed to allow for the decomposition (124).

Combining expressions (113c) and (114c), we find that h^i_{*j} must obey

$$\tilde{g}^{kl}(h^m_{*k|i|m|l} + h^m_{*i|k|m|l} - h^m_{*i|k|l|m}) = 0, \tag{116}$$

in addition to (111c). The relations (116) are, however, fulfilled identically for FLRW universes. This can easily be shown. First, we recall the well-known relation that the difference of the covariant derivatives $A^{i\cdots j}{}_{k\cdots l|p|q}$ and $A^{i\cdots j}{}_{k\cdots l|q|p}$ of an arbitrary tensor can be expressed in terms of the curvature and the tensor itself (Weinberg [35], Chapter 6, Section 5)

$$\begin{aligned} A^{i\cdots j}{}_{k\cdots l|p|q} - A^{i\cdots j}{}_{k\cdots l|q|p} = & \\ & + A^{i\cdots j}{}_{s\cdots l}\,{}^3R^s_{(0)kpq} + \cdots + A^{i\cdots j}{}_{k\cdots s}\,{}^3R^s_{(0)lpq} \\ & - A^{s\cdots j}{}_{k\cdots l}\,{}^3R^i_{(0)spq} - \cdots - A^{i\cdots s}{}_{k\cdots l}\,{}^3R^j_{(0)spq}, \end{aligned} \tag{117}$$

where ${}^3R^i_{(0)jkl}$ is the Riemann tensor for the spaces of constant time. At the right-hand side, there is a term with a plus sign for each lower index and a term with a minus sign for each upper index.

We apply this identity taking for A the second rank tensor h^i_{*j} to obtain

$$h^m_{*k|i|m} - h^m_{*k|m|i} = h^m_{*s}\,{}^3R^s_{(0)kim} - h^s_{*k}\,{}^3R^m_{(0)sim}. \tag{118}$$

Now note that $h^m_{*k|m|i}$ vanishes in view of (111c). Next, we take the covariant derivative of (118) with respect to x^l, and contract with $\tilde{g}^{kl}$

$$\tilde{g}^{kl} h^m_{*k|i|m|l} = \tilde{g}^{kl}(h^m_{*s}\,{}^3R^s_{(0)kim} - h^s_{*k}\,{}^3R^m_{(0)sim})_{|l}. \tag{119}$$

Next, using the expression which one has for the Riemann tensor of a maximally symmetric three-space,

$${}^3R^a_{(0)bcd} = k\,(\delta^a{}_c\tilde{g}_{bd} - \delta^a{}_d\tilde{g}_{bc})\,, \tag{120}$$

(where $k = 0, \pm 1$ is the curvature constant) we find

$$\tilde{g}^{kl} h^m_{*k|i|m|l} = 0, \tag{121}$$

i.e. the first term of (116) vanishes. The second and third term can similarly be expressed in the curvature

$$h^m_{*i|k|m|l} - h^m_{*i|k|l|m} = h^m_{*s|k}\,{}^3R^s_{(0)iml} + h^m_{*i|s}\,{}^3R^s_{(0)kml} - h^s_{*i|k}\,{}^3R^m_{(0)sml}, \tag{122}$$

where the general property (117) has been used. Upon substituting the Riemann tensor (120) and contracting with $\tilde{g}^{kl}$, we then arrive at

$$\tilde{g}^{kl}(h^m_{*i|k|m|l} - h^m_{*i|k|l|m}) = 0, \tag{123}$$

i.e. the second and third term of (116) together vanish. Hence, for FLRW universes, equation (116) is identically fulfilled. Consequently, the decomposition (111c) imposes no additional condition on the irreducible part h^i_{*j} of the perturbation $h^i{}_j$.

The three-vector $\boldsymbol{u}_{(1)}$ can be uniquely divided according to [37]

$$\boldsymbol{u}_{(1)} = \boldsymbol{u}_{(1)\parallel} + \boldsymbol{u}_{(1)\perp}, \tag{124}$$

where $\boldsymbol{u}_{(1)\parallel}$ is the *longitudinal* part of $\boldsymbol{u}_{(1)}$, with the properties

$$\tilde{\boldsymbol{\nabla}} \wedge (\boldsymbol{u}_{(1)\parallel}) = 0, \quad \tilde{\boldsymbol{\nabla}} \cdot \boldsymbol{u}_{(1)} = \tilde{\boldsymbol{\nabla}} \cdot \boldsymbol{u}_{(1)\parallel}, \tag{125}$$

and $\boldsymbol{u}_{(1)\perp}$ is the *transverse* part of $\boldsymbol{u}_{(1)}$, with the properties

$$\tilde{\boldsymbol{\nabla}} \cdot \boldsymbol{u}_{(1)\perp} = 0, \quad \tilde{\boldsymbol{\nabla}} \wedge \boldsymbol{u}_{(1)} = \tilde{\boldsymbol{\nabla}} \wedge \boldsymbol{u}_{(1)\perp}, \tag{126}$$

where the divergence of the vector $\boldsymbol{u}_{(1)}$ is defined by [see (47)]

$$\tilde{\boldsymbol{\nabla}} \cdot \boldsymbol{u}_{(1)} \equiv u^k_{(1)|k} = \vartheta_{(1)}, \tag{127}$$

and the rotation of the vector $\boldsymbol{u}_{(1)}$ is defined by

$$(\tilde{\boldsymbol{\nabla}} \wedge \boldsymbol{u}_{(1)})_i \equiv \epsilon_i{}^{jk} u_{(1)j|k} = \epsilon_i{}^{jk} u_{(1)j,k}, \tag{128}$$

where $\epsilon_i{}^{jk}$ is the Levi-Civita tensor with $\epsilon_1{}^{23} = +1$. In expression (128) we could replace the covariant derivative by the ordinary partial derivative because of the symmetry of $\Gamma^i{}_{jk}$.

Having decomposed the tensors $h^i{}_j$, ${}^3R^i_{(1)j}$ and $u^i_{(1)}$ in a scalar $\parallel$, a vector $\perp$ and a tensor part $*$, we can now decompose the set of equations (109) into three independent sets. The recipe is simple: all we have to do is to append a sub-index $\parallel$, $\perp$ or $*$ to the relevant tensorial quantities in equations (109). This will be the subject of the Sections VII A, VII B and VII C below.

A. Tensor Perturbations

We will show that tensor perturbations are not coupled to, i.e. do not give rise to, density perturbations. Upon substituting $h^i{}_j = h^i_{*j}$ and ${}^3R^i_{(1)j} = {}^3R^i_{(1)*j}$ into the perturbation equations (109) and using the properties (111c) and (113c), we find from equations (109a), (109b) and (109d)

$$\varepsilon_{(1)} = 0, \quad p_{(1)} = 0, \quad n_{(1)} = 0, \quad \boldsymbol{u}_{(1)} = 0, \tag{129}$$

where we have also used (77). With (129), equations (109e) and (109f) are identically satisfied. The only surviving equation is (109c), which now reads

$$\ddot{h}^i_{*j} + 3H\dot{h}^i_{*j} + 2\,{}^3R^i_{(1)*j} = 0, \tag{130}$$

where ${}^3R^i_{(1)*j}$ is given by (114c). Using (73), (74), (111c) and (129) it follows from (3) that

$$\varepsilon^{\mathrm{gi}}_{(1)} = 0, \quad n^{\mathrm{gi}}_{(1)} = 0, \tag{131}$$

so that tensor perturbations do not, in first-order, contribute to physical energy density and particle number density perturbations. Hence, the equations (130) do not play a role in this context, where we are interested in energy density and particle number density perturbations only.

The equations (130) have a wave equation like form with an extra term. Therefore, these tensor perturbations are sometimes called *gravitational waves*. The extra term $3H\dot{h}^i_{*j}$ in these equations is due to the expansion of the universe. The six components h^i_{*j} satisfy the four conditions (111c), leaving us with two independent functions h^i_{*j}. They are related to linearly and circularly polarized waves.

B. Vector Perturbations

We will show that, just like tensor perturbations, vector perturbations are not coupled to density perturbations. Upon replacing $h^i{}_j$ by $h^i_{\perp j}$ and ${}^3R^i_{(1)j}$ by ${}^3R^i_{(1)\perp j}$ in the perturbation equations (109), and using the expressions (111b) and (113b), we find from equation (109a) and the trace of equation (109c)

$$\varepsilon_{(1)} = 0, \quad p_{(1)} = 0, \quad n_{(1)} = 0, \tag{132}$$

where we have also used (77).

Since $h^i_{\perp j}$ is traceless and raising the index with $g^{ij}_{(0)}$ in equation (109b) we get

$$\dot{h}^{kj}_{\perp|k} + 2Hh^{kj}_{\perp|k} = 2\kappa(\varepsilon_{(0)} + p_{(0)})u^j_{(1)}, \tag{133}$$

where we have used (28) and (56). We now calculate the covariant derivative of equations (133) with respect to x^j, and use (115) to obtain

$$\tilde{\boldsymbol{\nabla}} \cdot \boldsymbol{u}_{(1)} = 0, \tag{134}$$

where we made use of the fact that the time derivative and the covariant derivative commute. With (124)–(126) we see that only the transverse part of $\boldsymbol{u}_{(1)}$, namely $\boldsymbol{u}_{(1)\perp}$, plays a role in vector perturbations. From (111b) and (132) it follows that the equations (109d) and (109f) are identically satisfied. The only surviving equations are (109b), (109c) and (109e), which now read

$$\dot{h}^k_{\perp i|k} = -2\kappa(\varepsilon_{(0)} + p_{(0)})u_{(1)\perp i}, \tag{135a}$$

$$\ddot{h}^i_{\perp j} + 3H\dot{h}^i_{\perp j} + 2\,{}^3R^i_{(1)\perp j} = 0, \tag{135b}$$

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\Big[(\varepsilon_{(0)} + p_{(0)})u^i_{(1)\perp}\Big] + 5H(\varepsilon_{(0)} + p_{(0)})u^i_{(1)\perp} = 0, \tag{135c}$$

where ${}^3R^i_{(1)\perp j}$ is given by (114b).

Using (73), (74), (132) and (134) we get from (3)

$$\varepsilon^{\mathrm{gi}}_{(1)} = 0, \quad n^{\mathrm{gi}}_{(1)} = 0, \tag{136}$$

implying that also vector perturbations do not, in first-order, contribute to physical energy density and particle number density perturbations. Hence, the equations (135) do not play a role when we are interested in energy density and particle number density perturbations, as we are here. Vector perturbations are also called *vortices*.

Since vector perturbations obey $\tilde{\boldsymbol{\nabla}} \cdot \boldsymbol{u}_{(1)\perp} = 0$, they have two degrees of freedom. As a consequence, the tensor $h^i_{\perp j}$ has also two degrees of freedom. These degrees of freedom are related to clockwise and counter-clockwise rotation of matter.

C. Scalar Perturbations

Differentiation of equations (109b) covariantly with respect to x^j we obtain

$$\dot{h}^k_{\parallel k|i|j} - \dot{h}^k_{\parallel i|k|j} = 2\kappa(\varepsilon_{(0)} + p_{(0)})u_{(1)\parallel i|j}. \tag{137}$$

Interchanging i and j in this equation, and subtracting the resulting equation from (137) we get

$$\dot{h}^k_{\parallel i|k|j} - \dot{h}^k_{\parallel j|k|i} = -2\kappa(\varepsilon_{(0)} + p_{(0)})(u_{(1)\parallel i|j} - u_{(1)\parallel j|i}), \tag{138}$$

where we have used that $\dot{h}^k_{\parallel k|i|j} = \dot{h}^k_{\parallel k|j|i}$. Using that $\tilde{\boldsymbol{\nabla}} \wedge \boldsymbol{u}_{(1)\parallel} = 0$, we find from (111a) that the function ζ must obey the equations

$$\dot{\zeta}^{|k}{}_{|i|k|j} - \dot{\zeta}^{|k}{}_{|j|k|i} = 0. \tag{139}$$

These equations are fulfilled identically in FLRW universes. This can be seen as follows. We first rewrite these equations by interchanging the covariant derivatives in the form

$$(\dot{\zeta}^{|k}{}_{|i|k|j} - \dot{\zeta}^{|k}{}_{|i|j|k}) - (\dot{\zeta}^{|k}{}_{|j|k|i} - \dot{\zeta}^{|k}{}_{|j|i|k}) + (\dot{\zeta}^{|k}{}_{|i|j} - \dot{\zeta}^{|k}{}_{|j|i})_{|k} = 0. \tag{140}$$

Next, we use expression (117) and substitute the Riemann tensor (120) into the resulting expression. Using that $\dot{\zeta}_{|i|j} = \dot{\zeta}_{|j|i}$, we find that the left-hand sides of the equations (140) vanish. As a consequence, the equations (139) are identities. Therefore, the decomposition (111a) imposes no additional condition on the irreducible part $h^i_{\parallel j}$ of the perturbation $h^i{}_j$.

We thus have shown that the system of first-order Einstein equations (109) is, for scalar perturbations, equivalent to the system

$$\text{Constraints:} \quad H\dot{h}^k_{\parallel k} + \tfrac{1}{2}\,{}^3\!R_{(1)\parallel} = -\kappa\varepsilon_{(1)}, \tag{141a}$$

$$\dot{h}^k_{\parallel k|i} - \dot{h}^k_{\parallel i|k} = 2\kappa(\varepsilon_{(0)} + p_{(0)})u_{(1)\parallel i}, \tag{141b}$$

$$\text{Evolution:} \quad \ddot{h}^i_{\parallel j} + 3H\dot{h}^i_{\parallel j} + \delta^i{}_j H\dot{h}^k_{\parallel k} + 2\,{}^3\!R^i_{(1)\parallel j} = -\kappa\delta^i{}_j(\varepsilon_{(1)} - p_{(1)}), \tag{141c}$$

$$\text{Conservation:} \quad \dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + (\varepsilon_{(0)} + p_{(0)})\theta_{(1)} = 0, \tag{141d}$$

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\Big[(\varepsilon_{(0)} + p_{(0)})u^i_{(1)\parallel}\Big] - g^{ik}_{(0)}p_{(1)|k} + 5H(\varepsilon_{(0)} + p_{(0)})u^i_{(1)\parallel} = 0, \tag{141e}$$

$$\dot{n}_{(1)} + 3Hn_{(1)} + n_{(0)}\theta_{(1)} = 0, \tag{141f}$$

where the local perturbations to the expansion, metric and the Ricci tensor are given by (73), (111a) and (114a) respectively. In the tensorial and vectorial case we found $\varepsilon_{(1)} = 0$ and $n_{(1)} = 0$, implying that $\varepsilon^{\mathrm{gi}}_{(1)} = 0$ and $n^{\mathrm{gi}}_{(1)} = 0$, which made the tensorial and vectorial equations irrelevant for our purpose. Such a conclusion cannot be drawn from the equations (141). Perturbations with $\varepsilon_{(1)} \neq 0$ and $n_{(1)} \neq 0$ are usually referred to as scalar perturbations.

Since the perturbation equations (141) contain only the components $h^i_{\parallel j}$, it follows that *relativistic* energy density and particle number density perturbations are characterized by *two* potentials, ϕ and ζ. In the next section we will rewrite the system (141) into a form which is suitable to study the evolution of the quantities (3).

VIII. FIRST-ORDER EQUATIONS FOR SCALAR PERTURBATIONS

The system of equations (141) can be further simplified by taking into account the decomposition (124). In the foregoing section we have shown that only the longitudinal part $\boldsymbol{u}_{(1)\parallel}$ of the three-vector $\boldsymbol{u}_{(1)}$ is coupled to density

perturbations. Using the properties (124)–(127) we can rewrite the scalar perturbation equations (141) in terms of quantities $\theta_{(1)}$, ${}^3\!R_{(1)\parallel}$, $\vartheta_{(1)}$, $\varepsilon_{(1)}$ and $n_{(1)}$, which are suitable to describe exclusively the scalar perturbations. The result is that the first-order quantities occurring in the definitions (3) do explicitly occur in the set of equations. A second important result is that the metric components $h^i_{\parallel j}$ occur, in the resulting equations, only in ${}^3\!R_{(1)\parallel}$. An evolution equation for this quantity follows from the $(0,i)$ perturbed constraint equations and will be derived below.

We now successively simplify all equations of the set (141) by replacing $\boldsymbol{u}_{(1)\parallel}$ by its divergence $\vartheta_{(1)}$, (74), and eliminating $\dot{h}^k_{\parallel k}$ with the help of (73) and using that $\dot{h}^k{}_k \equiv \dot{h}^k_{\parallel k}$, as follows from (110) and (111). For equation (141a) we find

$$2H(\theta_{(1)} - \vartheta_{(1)}) - \tfrac{1}{2}\,{}^3\!R_{(1)\parallel} = \kappa\varepsilon_{(1)}. \tag{142}$$

Thus the $(0,0)$-component of the constraint equations becomes an algebraic equation which relates the first-order quantities $\theta_{(1)}$, $\vartheta_{(1)}$, ${}^3\!R_{(1)\parallel}$ and $\varepsilon_{(1)}$. In (142), ${}^3\!R_{(1)\parallel}$ is given by

$${}^3\!R_{(1)\parallel} = g^{ij}_{(0)}(h^k_{\parallel k|i|j} - h^k_{\parallel i|j|k}) + \tfrac{1}{3}\,{}^3\!R_{(0)}h^k_{\parallel k}, \tag{143}$$

as follows from (87) with (111)–(113) and (115).

We will now derive an evolution equation for ${}^3\!R_{(1)\parallel}$ from equations (141b). Firstly, multiplying both sides of equations (141b) by $g^{ij}_{(0)}$ and taking the covariant divergence with respect to the index j we find

$$g^{ij}_{(0)}(\dot{h}^k_{\parallel k|i|j} - \dot{h}^k_{\parallel i|k|j}) = 2\kappa(\varepsilon_{(0)} + p_{(0)})\vartheta_{(1)}, \tag{144}$$

where we have also used (74). The left-hand side will turn up as a part of the time derivative of the curvature ${}^3\!R_{(1)\parallel}$. In fact, differentiating (143) with respect to ct and recalling that the connection coefficients $\Gamma^k_{(0)ij}$, (62), are independent of time, one gets

$${}^3\!\dot{R}_{(1)\parallel} = -2H\,{}^3\!R_{(1)\parallel} + g^{ij}_{(0)}(\dot{h}^k_{\parallel k|i|j} - \dot{h}^k_{\parallel i|k|j}) + \tfrac{1}{3}\,{}^3\!R_{(0)}\dot{h}^k_{\parallel k}, \tag{145}$$

where we have used (28), (56), (90) and

$$g^{ij}_{(0)}h^k_{\parallel i|j|k} = g^{ij}_{(0)}h^k_{\parallel i|k|j}, \tag{146}$$

which is a consequence of $g^{ij}_{(0)|k} = 0$ and the symmetry of $h^{ij}_{\parallel}$. Next, combining equation (144) with (145), and, finally, eliminating $\dot{h}^k_{\parallel k}$ with the help of (73), one arrives at

$${}^3\!\dot{R}_{(1)\parallel} + 2H\,{}^3\!R_{(1)\parallel} - 2\kappa(\varepsilon_{(0)} + p_{(0)})\vartheta_{(1)} + \tfrac{2}{3}\,{}^3\!R_{(0)}(\theta_{(1)} - \vartheta_{(1)}) = 0. \tag{147}$$

In this way we managed to recast the three $(0,i)$-components of the constraint equations in the form of one ordinary differential equation for the local perturbation, ${}^3\!R_{(1)\parallel}$, to the spatial curvature.

We now consider the dynamical equations (141c). Taking the trace of these equations and using (73) to eliminate the quantity $\dot{h}^k_{\parallel k}$, we arrive at

$$\dot{\theta}_{(1)} - \dot{\vartheta}_{(1)} + 6H(\theta_{(1)} - \vartheta_{(1)}) - {}^3\!R_{(1)\parallel} = \tfrac{3}{2}\kappa(\varepsilon_{(1)} - p_{(1)}). \tag{148}$$

Thus, for scalar perturbations, the three dynamical Einstein equations (141c) with $i = j$ reduce to one ordinary differential equation for the difference $\theta_{(1)} - \vartheta_{(1)}$. For $i \neq j$ the dynamical Einstein equations are not coupled to scalar perturbations. Therefore, these equations need not be considered.

Taking the covariant derivative of equation (141e) with respect to the metric $g_{(0)ij}$ and using (74), we get

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\Big[(\varepsilon_{(0)} + p_{(0)})\vartheta_{(1)}\Big] - g^{ik}_{(0)}p_{(1)|k|i} + 5H(\varepsilon_{(0)} + p_{(0)})\vartheta_{(1)} = 0, \tag{149}$$

where we have used that the operations of taking the time derivative and the covariant derivative commute, since the connection coefficients $\Gamma^k_{(0)ij}$, (62), are independent of time. With equation (92), we can rewrite equation (149) in the form

$$\dot{\vartheta}_{(1)} + H\left(2 - 3\frac{\dot{p}_{(0)}}{\dot{\varepsilon}_{(0)}}\right)\vartheta_{(1)} + \frac{1}{\varepsilon_{(0)} + p_{(0)}}\frac{\tilde{\nabla}^2 p_{(1)}}{a^2} = 0, \tag{150}$$

where $\tilde{\nabla}^2$ is the generalized Laplace operator which, for an arbitrary function $f(t, \boldsymbol{x})$ and with respect to an arbitrary three-dimensional metric $\tilde{g}^{ij}(\boldsymbol{x})$, is defined by

$$\tilde{\nabla}^2 f \equiv \tilde{g}^{ij} f_{|i|j} = \frac{1}{\sqrt{\det \tilde{g}}} \frac{\partial}{\partial x^i} \left(\tilde{g}^{ij} \sqrt{\det \tilde{g}} \, \frac{\partial f}{\partial x^j} \right). \tag{151}$$

Thus, the three first-order momentum conservation laws (141e) reduce to one ordinary differential equation for the divergence $\vartheta_{(1)}$.

Finally, the conservation laws (141d) and (141f) are already written in a suitable form. This concludes the derivation of the first-order equations for scalar perturbations.

We thus have shown that the system of equations (141) is equivalent to the system

$$\text{Constraint:} \quad 2H(\theta_{(1)} - \vartheta_{(1)}) - \tfrac{1}{2}\,{}^3\!R_{(1)\parallel} = \kappa\varepsilon_{(1)}, \tag{152a}$$

$$\text{Evolution:} \quad {}^3\!\dot{R}_{(1)\parallel} + 2H\,{}^3\!R_{(1)\parallel} - 2\kappa\varepsilon_{(0)}(1+w)\vartheta_{(1)} + \tfrac{2}{3}\,{}^3\!R_{(0)}(\theta_{(1)} - \vartheta_{(1)}) = 0, \tag{152b}$$

$$\dot{\theta}_{(1)} - \dot{\vartheta}_{(1)} + 6H(\theta_{(1)} - \vartheta_{(1)}) - {}^3\!R_{(1)\parallel} = \tfrac{3}{2}\kappa(\varepsilon_{(1)} - p_{(1)}), \tag{152c}$$

$$\text{Conservation:} \quad \dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + \varepsilon_{(0)}(1+w)\theta_{(1)} = 0, \tag{152d}$$

$$\dot{\vartheta}_{(1)} + H(2 - 3\beta^2)\vartheta_{(1)} + \frac{1}{\varepsilon_{(0)}(1+w)} \frac{\tilde{\nabla}^2 p_{(1)}}{a^2} = 0, \tag{152e}$$

$$\dot{n}_{(1)} + 3Hn_{(1)} + n_{(0)}\theta_{(1)} = 0. \tag{152f}$$

The quantities $\beta(t)$ and $w(t)$ occurring in equations (152) are defined by

$$\beta(t) \equiv \sqrt{\frac{\dot{p}_{(0)}(t)}{\dot{\varepsilon}_{(0)}(t)}}, \quad w(t) \equiv \frac{p_{(0)}(t)}{\varepsilon_{(0)}(t)}. \tag{153}$$

The algebraic equation (152a) and the five *ordinary* differential equations (152b)–(152f), is a system of six equations for the five quantities $\theta_{(1)}$, ${}^3\!R_{(1)}$, $\vartheta_{(1)}$, $\varepsilon_{(1)}$ and $n_{(1)}$ respectively. This system is, however, not over-determined: in exactly the same way as we have shown in Section VI C 2 for the background equations, it can be easily shown by differentiation of the constraint equation (152a) with respect to time that the general solution of the system (152a)–(152b) and (152d)–(152f) is also a solution of the dynamical equation (152c). Therefore, it is not needed to consider equation (152c) anymore.

IX. SUMMARY OF THE BASIC EQUATIONS

In this section we will show that there exist unique gauge-invariant density perturbations (3). To that end we summarize the background and first-order equations for scalar perturbations.

A. Zeroth-order Equations

The Einstein equations and conservation laws for the background FLRW universe are given by (57), (88), (90) and (92)–(93):

$$\text{Constraint:} \quad 3H^2 = \tfrac{1}{2}\,{}^3\!R_{(0)} + \kappa\varepsilon_{(0)} + \Lambda, \tag{154a}$$

$$\text{Evolution:} \quad {}^3\!\dot{R}_{(0)} = -2H\,{}^3\!R_{(0)}, \tag{154b}$$

$$\text{Conservation:} \quad \dot{\varepsilon}_{(0)} = -3H\varepsilon_{(0)}(1+w), \tag{154c}$$

$$\vartheta_{(0)} = 0, \tag{154d}$$

$$\dot{n}_{(0)} = -3Hn_{(0)}, \tag{154e}$$

where the initial value for (154b) is given by (91). As we have shown in Section VI C 2, the dynamical Einstein equation (89) is not needed. The set (154) consists of one algebraic and three differential equations with respect to time for the four unknown quantities $\varepsilon_{(0)}$, $n_{(0)}$, $\theta_{(0)} = 3H$ and ${}^3\!R_{(0)}$. The pressure $p_{(0)}(t)$ is related to the energy density $\varepsilon_{(0)}(t)$ and the particle number density $n_{(0)}(t)$ via the equation of state (59).

We rewrite the Friedmann equation (154a) by dividing both sides by $3H^2$ in the form

$$1 = \Omega_{\rm curv} + \Omega_{\rm bar} + \Omega_{\rm cdm} + \Omega_{\rm rad} + \Omega_\Lambda, \tag{155}$$

where $\Omega_{\rm curv}$, $\Omega_{\rm bar}$, $\Omega_{\rm cdm}$, $\Omega_{\rm rad}$ and Ω_Λ are the contributions due to curvature; baryonic (ordinary) matter; CDM; radiation and dark energy respectively. The time dependence of these contributions is given by

$$\Omega_{\rm curv}(t) \equiv -\frac{k}{a^2H^2}, \quad \Omega_{\rm bar}(t) + \Omega_{\rm cdm}(t) + \Omega_{\rm rad}(t) \equiv \frac{\kappa\varepsilon_{(0)}}{3H^2}, \quad \Omega_\Lambda(t) \equiv \frac{\Lambda}{3H^2}, \tag{156}$$

where we have used (65). This enables us to link the observations made with the WMAP satellite to our treatise on density perturbations. The present day values of the quantities in (155) are, for a ΛCDM universe (also referred to as the *concordance model*) [14–18], given by

$$\Omega_{\rm bar}(t_{\rm p}) = 0.0441, \quad \Omega_{\rm cdm}(t_{\rm p}) = 0.214, \quad \Omega_{\rm rad}(t_{\rm p}) = 0, \quad \Omega_\Lambda(t_{\rm p}) = 0.742. \tag{157}$$

With (155) we get

$$\Omega_{\rm curv}(t_{\rm p}) = 0.000. \tag{158}$$

Thus, WMAP-observations indicate that the universe is flat. Moreover, it follows from WMAP-observations that the present value of the Hubble function (54) is

$$\mathcal{H}(t_{\rm p}) = 71.9\,{\rm km/s/Mpc}. \tag{159}$$

Using that $1\,{\rm Mpc} = 3.0857\times10^{22}\,{\rm m}$, we get for the curvature parameter k and the cosmological constant Λ, using the observed values (157)–(159),

$$k = 0, \quad \Lambda = 1.34\times10^{-52}\,{\rm m}^{-2}, \tag{160}$$

respectively. The cosmological constant represents *dark energy*, also known as *quintessence*, a constant energy density filling space homogeneously. The existence of dark energy is postulated in order to explain recent observations that today the universe appears to be expanding at an *accelerating* rate. Since accelerated expansion takes place only at late times, we do not take into account Λ in our calculations of star formation in Sections XIII–XV.

In astrophysics and cosmology, CDM is hypothetical matter of unknown composition that does not emit or reflect enough electromagnetic radiation to be observed directly, but whose presence can be inferred from gravitational effects on visible matter. One of the main candidates for CDM are the so-called WIMPs (Weakly Interacting Massive Particles) with a mass of approximately 10–$10^3\,{\rm GeV}/c^2$. A recent estimate [38] yields a WIMP mass of approximately $70\,{\rm GeV}/c^2$. In comparison, the mass of a proton is $0.938\,{\rm GeV}/c^2$. Because of their large mass, WIMPs move relatively slow and are therefore cold. Since WIMPs do interact only through *weak nuclear force* with a range of approximately $10^{-17}\,{\rm m}$, they are dark and, as a consequence, invisible through electromagnetic observations. The only perceptible interaction with ordinary matter is through gravity. In the literature, CDM is therefore treated as 'dust,' i.e. a substance which interacts only through gravity with itself and ordinary matter.

B. First-order Equations

Since the evolution equation (152c) is not needed, as we have shown at the end of Section VIII, the first-order equations describing density perturbations are given by the set of one algebraic equation and four differential equations (152a)–(152b) and (152d)–(152f):

$$\text{Constraint:}\quad 2H(\theta_{(1)} - \vartheta_{(1)}) - \tfrac{1}{2}\,{}^3\!R_{(1)\|} = \kappa\varepsilon_{(1)}, \tag{161a}$$

$$\text{Evolution:}\quad {}^3\!\dot{R}_{(1)\|} + 2H\,{}^3\!R_{(1)\|} - 2\kappa\varepsilon_{(0)}(1+w)\vartheta_{(1)} + \tfrac{2}{3}\,{}^3\!R_{(0)}(\theta_{(1)} - \vartheta_{(1)}) = 0, \tag{161b}$$

$$\text{Conservation:}\quad \dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + \varepsilon_{(0)}(1+w)\theta_{(1)} = 0, \tag{161c}$$

$$\dot{\vartheta}_{(1)} + H(2-3\beta^2)\vartheta_{(1)} + \frac{1}{\varepsilon_{(0)}(1+w)}\frac{\tilde{\nabla}^2 p_{(1)}}{a^2} = 0, \tag{161d}$$

$$\dot{n}_{(1)} + 3Hn_{(1)} + n_{(0)}\theta_{(1)} = 0, \tag{161e}$$

for the five unknown functions $\varepsilon_{(1)}$, $n_{(1)}$, $\vartheta_{(1)}$, ${}^3\!R_{(1)\|}$ and $\theta_{(1)}$ respectively. The first-order perturbation to the pressure is given by the perturbed equation of state (77).

From their derivation it follows that the set of equations (161) is equivalent to the system of equations (141). The metric is in the set (161) contained in only one quantity, namely the local first-order perturbation ${}^3R_{(1)\parallel}$ to the global spatial curvature. Using the sets of equations (154) and (161), we show in Section XII that our perturbation theory yields the Newtonian theory of gravity in the non-relativistic limit of an expanding universe with (3) as the key quantities.

C. Unique Gauge-invariant Density Perturbations

The background equations (154) and first-order equations (161) are now rewritten in such a form that we can draw an important conclusion. Firstly, we observe that equations (161) are the first-order counterparts of the background equations (154). Combined, these two sets of equations describe the background quantities and their corresponding first-order quantities:

$$(\varepsilon_{(0)}, \varepsilon_{(1)}), \quad (n_{(0)}, n_{(1)}), \quad (\vartheta_{(0)} = 0, \vartheta_{(1)}), \quad ({}^3R_{(0)}, {}^3R_{(1)\parallel}), \quad (\theta_{(0)} = 3H, \theta_{(1)}). \tag{162}$$

Secondly, we remark that the quantities ϑ and ${}^3R$ are not scalars of space-time, since their first-order perturbations $\vartheta_{(1)}$ and ${}^3R_{(1)\parallel}$ transform according to (C7) and (C9) respectively. In contrast, the first-order perturbations $\varepsilon_{(1)}$, $n_{(1)}$ and $\theta_{(1)}$ of the scalars (20) transform according to (15). Thus, *only three independent scalars* ε, n and θ, (20), play a role in a density perturbation theory. Consequently, the only non-trivial gauge-invariant combinations which can be constructed from these scalars and their first-order perturbations are the combinations (3). We thus have shown that there exist *unique* gauge-invariant quantities $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$. In Section XII we show that a perturbation theory based on these quantities yields the Newtonian theory of gravity in the non-relativistic limit of an expanding universe.

By switching from the variables $\varepsilon_{(1)}$, $n_{(1)}$, $\vartheta_{(1)}$, ${}^3R_{(1)\parallel}$ and $\theta_{(1)}$ to the variables $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$, we will arrive at a set of equations for $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ with a unique, i.e. gauge-invariant solution. This will be the subject of Section XI. First, we derive some auxiliary expressions related to the entropy, pressure and temperature.

X. ENTROPY, PRESSURE, TEMPERATURE AND DIABATIC PERTURBATIONS

In order to study the evolution of density perturbations, we need the laws of thermodynamics. In this section we rewrite the combined First and Second Laws of thermodynamics in terms of the gauge-invariant quantities (3) and the gauge-invariant entropy per particle $s^{\mathrm{gi}}_{(1)}$.

A. Gauge-invariant Entropy Perturbations

Consider a gas of N particles with volume V. Let μ be the thermodynamic —or chemical— potential, p the pressure and S its entropy. Then the internal energy E is given by the relation

$$E = TS - pV + \mu N. \tag{163}$$

In terms of the energy per particle $e \equiv E/N$, the entropy per particle $s \equiv S/N$ and the particle number density $n \equiv N/V$ this relation reads

$$e = Ts - pn^{-1} + \mu, \tag{164}$$

implying that

$$\mathrm{d}e = T\mathrm{d}s + s\mathrm{d}T - n^{-1}\mathrm{d}p - p\mathrm{d}n^{-1} + \mathrm{d}\mu. \tag{165}$$

From (163) we also find

$$\mathrm{d}E = T\mathrm{d}S + S\mathrm{d}T - V\mathrm{d}p - p\mathrm{d}V + \mu\mathrm{d}N + N\mathrm{d}\mu. \tag{166}$$

The combined First and Second Laws of thermodynamics reads

$$\mathrm{d}E = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N. \tag{167}$$

From (166) and (167) we find after division by N

$$s\mathrm{d}T - n^{-1}\mathrm{d}p + \mathrm{d}\mu = 0. \tag{168}$$

This relation enables us to eliminate $\mathrm{d}\mu$ from (165). We so find

$$T\mathrm{d}s = \mathrm{d}e + p\mathrm{d}n^{-1}. \tag{169}$$

This thermodynamic relation is independent of N and μ. In terms of the energy density defined as $\varepsilon \equiv ne$, we so find, finally,

$$T\mathrm{d}s = \mathrm{d}\left(\frac{\varepsilon}{n}\right) + p\mathrm{d}\left(\frac{1}{n}\right). \tag{170}$$

This is the relation we shall use in the following.

The thermodynamic relation (170) is true for a system in thermodynamic equilibrium. For a non-equilibrium system that is 'not too far' from equilibrium, the equation (170) may be replaced by

$$T\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\varepsilon}{n}\right) + p\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{n}\right), \tag{171}$$

where $\mathrm{d}/\mathrm{d}t$ is the time derivative in a local co-moving Lorentz system. Now, using $\varepsilon = \varepsilon_{(0)} + \varepsilon_{(1)}$, $s = s_{(0)} + s_{(1)}$, $p = p_{(0)} + p_{(1)}$ and $n = n_{(0)} + n_{(1)}$, we find from equation (171)

$$T_{(0)}\frac{\mathrm{d}s_{(0)}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\varepsilon_{(0)}}{n_{(0)}}\right) + p_{(0)}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{n_{(0)}}\right), \tag{172}$$

where we neglected time derivatives of first-order quantities. With the help of equations (154c), (154e) and (153) we find that the right-hand side of equation (172) vanishes. Hence, $\dot{s}_{(0)} = 0$, implying that, in zeroth-order, the expansion takes place without generating entropy: $s_{(0)}$ is constant in time. Hence, in view of (15), which is valid for any scalar, and $\dot{s}_{(0)} = 0$, the first-order perturbation $s_{(1)}$ is automatically a gauge-invariant quantity, i.e. $\hat{s}_{(1)} = s_{(1)}$, in contrast to $\varepsilon_{(1)}$ and $n_{(1)}$, which had to be redefined according to expressions (3). Apparently, the entropy per particle $s_{(1)}$ is such a combination of $\varepsilon_{(1)}$ and $n_{(1)}$ that it need not be redefined. This can be made explicit by noting that in the linear approximation we are considering, the combined First and Second Laws of thermodynamics (170) should hold for zeroth-order and first-order quantities separately. In particular, equation (170) implies

$$T_{(0)}s_{(1)} = \frac{1}{n_{(0)}}\left(\varepsilon_{(1)} - \frac{\varepsilon_{(0)} + p_{(0)}}{n_{(0)}}n_{(1)}\right), \tag{173}$$

where we neglected products of differentials and first-order quantities, and where we replaced $\mathrm{d}\varepsilon$ and $\mathrm{d}n$ by $\varepsilon_{(1)}$ and $n_{(1)}$ respectively. We now note that the linear combination in the right-hand side of equation (173) has the property

$$\varepsilon_{(1)} - \frac{\varepsilon_{(0)} + p_{(0)}}{n_{(0)}}n_{(1)} = \varepsilon^{\mathrm{gi}}_{(1)} - \frac{\varepsilon_{(0)} + p_{(0)}}{n_{(0)}}n^{\mathrm{gi}}_{(1)}, \tag{174}$$

as may immediately be verified with the help of (3) and the equations (154c) and (154e). The right-hand side of expression (174) being gauge-invariant, the left-hand side must be gauge-invariant. This observation makes explicit the gauge-invariance of the first-order approximation to the entropy per particle, $s_{(1)}$. In order to stress the gauge-invariance of the correction $s_{(1)}$ to the (constant) entropy per particle, $s_{(0)}$, we will write $s^{\mathrm{gi}}_{(1)}$, rather than $s_{(1)}$. From (173), (174) and $s^{\mathrm{gi}}_{(1)} \equiv s_{(1)}$ we then get

$$T_{(0)}s^{\mathrm{gi}}_{(1)} = \frac{1}{n_{(0)}}\left(\varepsilon^{\mathrm{gi}}_{(1)} - \frac{\varepsilon_{(0)}(1+w)}{n_{(0)}}n^{\mathrm{gi}}_{(1)}\right), \tag{175}$$

where w is the quotient of zeroth-order pressure and zeroth-order energy density defined by (153). With (175) we have rewritten the thermodynamic law (167) in terms of the quantities $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$.

We rewrite equation (175) in the form

$$T_{(0)}s^{\mathrm{gi}}_{(1)} = -\frac{\varepsilon_{(0)}(1+w)}{n^2_{(0)}}\sigma^{\mathrm{gi}}_{(1)}, \tag{176}$$

where the gauge-invariant, entropy related quantity $\sigma^{\mathrm{gi}}_{(1)}$ is given by

$$\sigma^{\mathrm{gi}}_{(1)} \equiv n^{\mathrm{gi}}_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)}\varepsilon^{\mathrm{gi}}_{(1)}. \tag{177}$$

The quantity $\sigma^{\mathrm{gi}}_{(1)}$ occurs as the source term in the evolution equations (197) below.

B. Gauge-invariant Pressure Perturbations

We will now derive a gauge-invariant expression for the physical pressure perturbations. To that end, we first calculate the time derivative of the background pressure. From the equation of state (59) we have

$$\dot{p}_{(0)} = p_n \dot{n}_{(0)} + p_\varepsilon \dot{\varepsilon}_{(0)}, \tag{178}$$

where p_ε and p_n are the partial derivatives given by expressions (78) and (79). Multiplying both sides of this expression by $\theta_{(1)}/\dot{\theta}_{(0)}$ and subtracting the result from (77) we get

$$p_{(1)} - \frac{\dot{p}_{(0)}}{\dot{\theta}_{(0)}}\theta_{(1)} = p_n n^{\mathrm{gi}}_{(1)} + p_\varepsilon \varepsilon^{\mathrm{gi}}_{(1)}, \tag{179}$$

where we have used (3) to rewrite the right-hand side. Since p_n and p_ε depend on the background quantities $\varepsilon_{(0)}$ and $n_{(0)}$ only, the right-hand side is gauge-invariant. Hence, the quantity $p^{\mathrm{gi}}_{(1)}$ defined by

$$p^{\mathrm{gi}}_{(1)} \equiv p_{(1)} - \frac{\dot{p}_{(0)}}{\dot{\theta}_{(0)}}\theta_{(1)}, \tag{180}$$

is gauge-invariant. We thus obtain the gauge-invariant counterpart of (77)

$$p^{\mathrm{gi}}_{(1)} = p_\varepsilon \varepsilon^{\mathrm{gi}}_{(1)} + p_n n^{\mathrm{gi}}_{(1)}. \tag{181}$$

We will now rewrite this expression in a slightly different form. From (153) and (178) we obtain $\beta^2 = p_\varepsilon + p_n(\dot{n}_{(0)}/\dot{\varepsilon}_{(0)})$. Using equations (154c) and (154e) we find

$$\beta^2 = p_\varepsilon + \frac{n_{(0)} p_n}{\varepsilon_{(0)}(1+w)}. \tag{182}$$

With this expression and (177) and (182) we can rewrite the pressure perturbation (181) as

$$p^{\mathrm{gi}}_{(1)} = \beta^2 \varepsilon^{\mathrm{gi}}_{(1)} + p_n \sigma^{\mathrm{gi}}_{(1)}. \tag{183}$$

We thus have expressed the pressure perturbation $p^{\mathrm{gi}}_{(1)}$ in terms of the energy density perturbation $\varepsilon^{\mathrm{gi}}_{(1)}$ and the entropy related quantity $\sigma^{\mathrm{gi}}_{(1)}$ rather than $\varepsilon^{\mathrm{gi}}_{(1)}$ and the particle number density perturbation $n^{\mathrm{gi}}_{(1)}$.

Expression (183) can be rewritten into an equivalent expression containing the entropy perturbation $s^{\mathrm{gi}}_{(1)}$ explicitly. For p_n we find

$$p_n \equiv \left(\frac{\partial p}{\partial n}\right)_\varepsilon = \left(\frac{\partial p}{\partial s}\right)_\varepsilon \left(\frac{\partial s}{\partial n}\right)_\varepsilon = -\frac{\varepsilon_{(0)}(1+w)}{n^2_{(0)} T_{(0)}} p_s, \quad p_s \equiv \left(\frac{\partial p}{\partial s}\right)_\varepsilon, \tag{184}$$

where we have used (175). Combining (176) and (183) we arrive at

$$p^{\mathrm{gi}}_{(1)} = \beta^2 \varepsilon^{\mathrm{gi}}_{(1)} + p_s s^{\mathrm{gi}}_{(1)}. \tag{185}$$

Substituting (184) into (182), we get

$$\beta^2 = p_\varepsilon - \frac{p_s}{n_{(0)} T_{(0)}}. \tag{186}$$

Expressions (182)–(183) refer to an equation of state $p = p(n, \varepsilon)$, whereas expressions (185)–(186) refer to the equivalent equation of state $p = p(s, \varepsilon)$.

C. Gauge-invariant Temperature Perturbations

Finally, we will derive an expression for the gauge-invariant temperature perturbation $T^{\mathrm{gi}}_{(1)}$ with the help of (A2a). For the time derivative of the energy density $\varepsilon_{(0)}(n_{(0)}, T_{(0)})$ we have

$$\dot{\varepsilon}_{(0)} = \left(\frac{\partial \varepsilon}{\partial n}\right)_T \dot{n}_{(0)} + \left(\frac{\partial \varepsilon}{\partial T}\right)_n \dot{T}_{(0)}. \tag{187}$$

Replacing the infinitesimal quantities in (A2a) by perturbations, we find

$$\varepsilon_{(1)} = \left(\frac{\partial\varepsilon}{\partial n}\right)_T n_{(1)} + \left(\frac{\partial\varepsilon}{\partial T}\right)_n T_{(1)}. \tag{188}$$

Multiplying both sides of (187) by $\theta_{(1)}/\dot{\theta}_{(0)}$ and subtracting the result from (188) we get

$$\varepsilon^{\mathrm{gi}}_{(1)} = \left(\frac{\partial\varepsilon}{\partial n}\right)_T n^{\mathrm{gi}}_{(1)} + \left(\frac{\partial\varepsilon}{\partial T}\right)_n \left(T_{(1)} - \frac{\dot{T}_{(0)}}{\dot{\theta}_{(0)}}\theta_{(1)}\right), \tag{189}$$

where we have used (3). Hence, the quantity

$$T^{\mathrm{gi}}_{(1)} \equiv T_{(1)} - \frac{\dot{T}_{(0)}}{\dot{\theta}_{(0)}}\theta_{(1)}, \tag{190}$$

is gauge-invariant. Thus, (189) can be written as

$$\varepsilon^{\mathrm{gi}}_{(1)} = \left(\frac{\partial\varepsilon}{\partial n}\right)_T n^{\mathrm{gi}}_{(1)} + \left(\frac{\partial\varepsilon}{\partial T}\right)_n T^{\mathrm{gi}}_{(1)}, \tag{191}$$

implying that $T^{\mathrm{gi}}_{(1)}$ can be interpreted as the gauge-invariant temperature perturbation. We thus have expressed the perturbation in the absolute temperature as a function of the perturbations in the energy density and particle number density for a given equation of state of the form $\varepsilon = \varepsilon(n, T)$ and $p = p(n, T)$. This expression will be used in Section XI to derive an expression for the fluctuations in the background temperature, δ_T, a measurable quantity.

Finally, we give the evolution equation for the background temperature $T_{(0)}(t)$. From (187) it follows that

$$\dot{T}_{(0)} = \frac{-3H\left[\varepsilon_{(0)}(1+w) - \left(\dfrac{\partial\varepsilon}{\partial n}\right)_T n_{(0)}\right]}{\left(\dfrac{\partial\varepsilon}{\partial T}\right)_n}, \tag{192}$$

where we have used equations (154c) and (154e). This equation will be used to follow the time development of the background temperature once $\varepsilon_{(0)}(t)$ and $n_{(0)}(t)$ are found from the zeroth-order Einstein equations.

D. Diabatic Perturbations in a FLRW Universe

In Section X A we have shown that the universe expands adiabatically. In this section we investigate under which conditions *local* density perturbations are adiabatic.

By definition, an *adiabatic*, or *isocaloric* process is a thermodynamic process in which no heat is transferred to or from the working fluid, i.e. it is a process for which $\delta Q = 0$. For a *reversible* process we have $\delta Q \equiv T_{(0)} s^{\mathrm{gi}}_{(1)}$. Hence, reversible and adiabatic processes are characterized by $T_{(0)} s^{\mathrm{gi}}_{(1)} = 0$. From expression (175) it follows that for adiabatic perturbations we have

$$n_{(0)}\varepsilon^{\mathrm{gi}}_{(1)} - \varepsilon_{(0)}(1+w)n^{\mathrm{gi}}_{(1)} = 0. \tag{193}$$

Using the background conservation laws (154c) and (154e), we get the adiabatic condition for density perturbations in a FLRW universe

$$\dot{n}_{(0)}\varepsilon^{\mathrm{gi}}_{(1)} - \dot{\varepsilon}_{(0)} n^{\mathrm{gi}}_{(1)} = 0. \tag{194}$$

In a non-static universe we have $\dot{\varepsilon}_{(0)} \neq 0$ and $\dot{n}_{(0)} \neq 0$. In this case, equation (194) is fulfilled if and only if the energy density is a function of the particle number density *only*, i.e. if $\varepsilon = \varepsilon(n)$. In particular, density perturbations in a perfect pressureless fluid with $\varepsilon = nmc^2$ are adiabatic. This is the case in the non-relativistic limit $v/c \to 0$ of an expanding FLRW universe, see Section XII. In all other cases, $\varepsilon = \varepsilon(n, T)$ and $p = p(n, T)$, local density perturbations evolve *diabatically*.

XI. MANIFESTLY GAUGE-INVARIANT FIRST-ORDER EQUATIONS

The five perturbation equations (161) form a set of five equations for the five unknown quantities $\varepsilon_{(1)}$, $n_{(1)}$, $\vartheta_{(1)}$, ${}^3R_{(1)\parallel}$ and $\theta_{(1)}$. This system of equations can be further reduced in the following way. As has been explained in Section IV B, our perturbation theory yields automatically $\theta^{\rm gi}_{(1)} = 0$, (25). As a consequence, we do not need the gauge dependent quantity $\theta_{(1)}$. Eliminating the quantity $\theta_{(1)}$ from equations (161) with the help of equation (161a), we arrive at the set of four first-order differential equations

$$\dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + \varepsilon_{(0)}(1+w)\Big[\vartheta_{(1)} + \frac{1}{2H}\left(\kappa\varepsilon_{(1)} + \tfrac{1}{2}\,{}^3R_{(1)\parallel}\right)\Big] = 0, \tag{195a}$$

$$\dot{n}_{(1)} + 3Hn_{(1)} + n_{(0)}\Big[\vartheta_{(1)} + \frac{1}{2H}\left(\kappa\varepsilon_{(1)} + \tfrac{1}{2}\,{}^3R_{(1)\parallel}\right)\Big] = 0, \tag{195b}$$

$$\dot{\vartheta}_{(1)} + H(2-3\beta^2)\vartheta_{(1)} + \frac{1}{\varepsilon_{(0)}(1+w)}\frac{\tilde{\nabla}^2 p_{(1)}}{a^2} = 0, \tag{195c}$$

$${}^3\dot{R}_{(1)\parallel} + 2H\,{}^3R_{(1)\parallel} - 2\kappa\varepsilon_{(0)}(1+w)\vartheta_{(1)} + \frac{{}^3R_{(0)}}{3H}\left(\kappa\varepsilon_{(1)} + \tfrac{1}{2}\,{}^3R_{(1)\parallel}\right) = 0, \tag{195d}$$

for the four quantities $\varepsilon_{(1)}$, $n_{(1)}$, $\vartheta_{(1)}$ and ${}^3R_{(1)\parallel}$. From their derivation it follows that the system of equations (195) is, for scalar perturbations, equivalent to the full set (109) of first-order Einstein equations.

The system of equations (195) is now cast in a suitable form to arrive at a system of manifestly gauge-invariant equations for the physical quantities $\varepsilon^{\rm gi}_{(1)}$ and $n^{\rm gi}_{(1)}$, since we then can immediately calculate these quantities. Indeed, eliminating the quantity $\theta_{(1)}$ from equations (3) with the help of equation (161a), and using the background equations (154) to eliminate the time derivatives $\dot{\varepsilon}_{(0)}$, $\dot{n}_{(0)}$ and $\dot{\theta}_{(0)} = 3\dot{H}$, we get

$$\varepsilon^{\rm gi}_{(1)} = \frac{\varepsilon_{(1)}\,{}^3R_{(0)} - 3\varepsilon_{(0)}(1+w)(2H\vartheta_{(1)} + \frac{1}{2}\,{}^3R_{(1)\parallel})}{{}^3R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}, \tag{196a}$$

$$n^{\rm gi}_{(1)} = n_{(1)} - \frac{3n_{(0)}(\kappa\varepsilon_{(1)} + 2H\vartheta_{(1)} + \frac{1}{2}\,{}^3R_{(1)\parallel})}{{}^3R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}. \tag{196b}$$

The quantities $\varepsilon^{\rm gi}_{(1)}$ and $n^{\rm gi}_{(1)}$ are now completely determined by the system of background equations (154) and the first-order equations (195).

A. Evolution Equations for Density Perturbations

Instead of calculating $\varepsilon^{\rm gi}_{(1)}$ and $n^{\rm gi}_{(1)}$ in the way described above, we proceed by first making explicit the gauge-invariance of the theory. To that end, we rewrite the system of four differential equations (195) for the gauge dependent variables $\varepsilon_{(1)}$, $n_{(1)}$, $\vartheta_{(1)}$ and ${}^3R_{(1)\parallel}$ into a system of equations for the gauge-invariant variables $\varepsilon^{\rm gi}_{(1)}$ and $n^{\rm gi}_{(1)}$. It is, however, of convenience to use the entropy related perturbation $\sigma^{\rm gi}_{(1)}$, defined by (177), rather than the particle number density perturbation $n^{\rm gi}_{(1)}$. The result is

$$\ddot{\varepsilon}^{\rm gi}_{(1)} + a_1\dot{\varepsilon}^{\rm gi}_{(1)} + a_2\varepsilon^{\rm gi}_{(1)} = a_3\sigma^{\rm gi}_{(1)}, \tag{197a}$$

$$\dot{\sigma}^{\rm gi}_{(1)} = -3H\left(1 - \frac{n_{(0)}p_n}{\varepsilon_{(0)}(1+w)}\right)\sigma^{\rm gi}_{(1)}. \tag{197b}$$

The derivation of these equations is given in detail in Appendices B 1 and B 2. The coefficients a_1, a_2 and a_3 occurring in equations (197) are given by

$$a_1 = \frac{\kappa\varepsilon_{(0)}(1+w)}{H} - 2\frac{\dot{\beta}}{\beta} + H(4-3\beta^2) + {}^3R_{(0)}\left(\frac{1}{3H} + \frac{2H(1+3\beta^2)}{{}^3R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}\right), \tag{198a}$$

$$a_2 = \kappa\varepsilon_{(0)}(1+w) - 4H\frac{\dot{\beta}}{\beta} + 2H^2(2-3\beta^2) + {}^3R_{(0)}\left(\frac{1}{2} + \frac{5H^2(1+3\beta^2) - 2H\dfrac{\dot{\beta}}{\beta}}{{}^3R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}\right) - \beta^2\left(\frac{\tilde{\nabla}^2}{a^2} - \tfrac{1}{2}\,{}^3R_{(0)}\right), \tag{198b}$$

$$a_3 = \left\{\frac{-18H^2}{{}^3R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}\left[\varepsilon_{(0)}p_{\varepsilon n}(1+w) + \frac{2p_n}{3H}\frac{\dot{\beta}}{\beta} - \beta^2 p_n + p_\varepsilon p_n + n_{(0)}p_{nn}\right] + p_n\right\}\left(\frac{\tilde{\nabla}^2}{a^2} - \tfrac{1}{2}\,{}^3R_{(0)}\right), \tag{198c}$$

where the functions $\beta(t)$ and $w(t)$ are given by (153). In the derivation of the above results, we used equations (154). The abbreviations p_n and p_ε are given by (78). Furthermore, we used the abbreviations

$$p_{nn} \equiv \frac{\partial^2 p}{\partial n^2}, \quad p_{\varepsilon n} \equiv \frac{\partial^2 p}{\partial \varepsilon\, \partial n}. \tag{199}$$

The equations (197) contain only gauge-invariant quantities and the coefficients are scalar functions. Thus, these equations are *manifestly gauge-invariant.* In contrast, the equations (195), being linear Einstein equations and conservation laws, are themselves gauge-invariant, but their solutions are gauge dependent, see Appendix C for a detailed explanation.

The equations (197) are equivalent to one equation of the third-order, whereas the four first-order equations (195) are equivalent to one equation of the fourth-order. This observation reflects the fact that the solutions of the four first-order equations (195) are gauge dependent, while the solutions $\varepsilon^{\rm gi}_{(1)}$ and $\sigma^{\rm gi}_{(1)}$ of equations (197) are gauge-invariant. One 'degree of freedom,' say, the gauge function $\psi(\boldsymbol{x})$ has disappeared from the scene completely.

The equations (197) constitute the main result of this article. In view of (177), they essentially are two differential equations for the perturbations $\varepsilon^{\rm gi}_{(1)}$ and $n^{\rm gi}_{(1)}$ to the energy density $\varepsilon_{(0)}(t)$ and the particle number density $n_{(0)}(t)$ respectively, for FLRW universes with $k=-1,0,+1$. They describe the evolution of the energy density perturbation $\varepsilon^{\rm gi}_{(1)}$ and the particle number density perturbation $n^{\rm gi}_{(1)}$ for FLRW universes filled with a fluid which is described by an equation of state of the form $p=p(n,\varepsilon)$, the precise form of which is left unspecified.

B. Evolution Equations for Contrast Functions

In the study of the evolution of density perturbations it is of convenience to use a quantity which measures the perturbation to the density relative to the background densities. To that end we define the gauge-invariant contrast functions δ_ε and δ_n by

$$\delta_\varepsilon(t,\boldsymbol{x}) \equiv \frac{\varepsilon^{\rm gi}_{(1)}(t,\boldsymbol{x})}{\varepsilon_{(0)}(t)}, \quad \delta_n(t,\boldsymbol{x}) \equiv \frac{n^{\rm gi}_{(1)}(t,\boldsymbol{x})}{n_{(0)}(t)}. \tag{200}$$

Using these quantities, equations (197) can be rewritten as (Appendix B 3)

$$\ddot{\delta}_\varepsilon + b_1\dot{\delta}_\varepsilon + b_2\delta_\varepsilon = b_3\left(\delta_n - \frac{\delta_\varepsilon}{1+w}\right), \tag{201a}$$

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta_n - \frac{\delta_\varepsilon}{1+w}\right) = \frac{3Hn_{(0)}p_n}{\varepsilon_{(0)}(1+w)}\left(\delta_n - \frac{\delta_\varepsilon}{1+w}\right), \tag{201b}$$

where the coefficients b_1, b_2 and b_3 are given by

$$b_1 = \frac{\kappa\varepsilon_{(0)}(1+w)}{H} - 2\frac{\dot\beta}{\beta} - H(2+6w+3\beta^2) + {}^3\!R_{(0)}\left(\frac{1}{3H} + \frac{2H(1+3\beta^2)}{{}^3\!R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}\right), \tag{202a}$$

$$\begin{aligned} b_2 = {} & -\tfrac{1}{2}\kappa\varepsilon_{(0)}(1+w)(1+3w) + H^2\left(1-3w+6\beta^2(2+3w)\right) \\ & + 6H\frac{\dot\beta}{\beta}\left(w + \frac{\kappa\varepsilon_{(0)}(1+w)}{{}^3\!R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}\right) - {}^3\!R_{(0)}\left(\tfrac{1}{2}w + \frac{H^2(1+6w)(1+3\beta^2)}{{}^3\!R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}\right) - \beta^2\left(\frac{\tilde\nabla^2}{a^2} - \tfrac{1}{2}\,{}^3\!R_{(0)}\right), \end{aligned} \tag{202b}$$

$$b_3 = \left\{\frac{-18H^2}{{}^3\!R_{(0)} + 3\kappa\varepsilon_{(0)}(1+w)}\left[\varepsilon_{(0)}p_{\varepsilon n}(1+w) + \frac{2p_n}{3H}\frac{\dot\beta}{\beta} - \beta^2 p_n + p_\varepsilon p_n + n_{(0)}p_{nn}\right] + p_n\right\}\frac{n_{(0)}}{\varepsilon_{(0)}}\left(\frac{\tilde\nabla^2}{a^2} - \tfrac{1}{2}\,{}^3\!R_{(0)}\right). \tag{202c}$$

In Sections XIII–XV we use the equations (201) to study the evolution of small energy density perturbations and particle number perturbations in FLRW universes.

The combined First and Second Laws of thermodynamics (175) reads, in terms of the contrast functions (200),

$$T_{(0)}s^{\rm gi}_{(1)} = -\frac{\varepsilon_{(0)}(1+w)}{n_{(0)}}\left(\delta_n - \frac{\delta_\varepsilon}{1+w}\right) = -\frac{\varepsilon_{(0)}}{n_{(0)}}(\delta_n - \delta_\varepsilon) - \frac{p_{(0)}}{n_{(0)}}\delta_n, \tag{203}$$

where $s^{\rm gi}_{(1)}$ is the entropy per particle.

Finally, we define the relative temperature perturbation δ_T and the relative pressure perturbation δ_p by

$$\delta_T(t,\boldsymbol{x}) \equiv \frac{T^{\mathrm{gi}}_{(1)}(t,\boldsymbol{x})}{T_{(0)}(t)}, \quad \delta_p(t,\boldsymbol{x}) \equiv \frac{p^{\mathrm{gi}}_{(1)}(t,\boldsymbol{x})}{p_{(0)}(t)}. \tag{204}$$

Using the expressions (191), (200) and (204) we arrive at the relative temperature perturbation

$$\delta_T = \frac{\varepsilon_{(0)}\delta_\varepsilon - \left(\dfrac{\partial\varepsilon}{\partial n}\right)_T n_{(0)}\delta_n}{T_{(0)}\left(\dfrac{\partial\varepsilon}{\partial T}\right)_n}. \tag{205}$$

The relative pressure perturbation follows directly from (181). We get

$$\delta_p = \frac{\varepsilon_{(0)}}{p_{(0)}}\left(\frac{\partial p}{\partial\varepsilon}\right)_n \delta_\varepsilon + \frac{n_{(0)}}{p_{(0)}}\left(\frac{\partial p}{\partial n}\right)_\varepsilon \delta_n. \tag{206}$$

We thus have found the relative temperature and the relative pressure perturbations as functions of the relative perturbations in the energy density and particle number density for an equation of state of the form $\varepsilon = \varepsilon(n,T)$ and $p = p(n,T)$ (see Appendix A). The quantity $\delta_T(t,\boldsymbol{x})$ is a measurable quantity in the cosmic background radiation.

XII. NON-RELATIVISTIC LIMIT IN AN EXPANDING FLRW UNIVERSE

In Section IX C we have shown that there exist only two gauge-invariant quantities $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$, which could be the real energy density and particle number density perturbations. In this section we show that in the non-relativistic limit $v/c \to 0$ the quantities $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ become equal to their Newtonian counterparts. This implies that $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ are indeed the local perturbations to the energy density and particle number density perturbations.

It is well known that if the gravitational field is weak and velocities are small with respect to the velocity of light ($v/c \to 0$), the system of Einstein equations and conservation laws may reduce to the single field equation of the Newtonian theory of gravity, namely the Poisson equation (229). In the first-order perturbation theory developed in this article, the gravitational field is already weak, so that, at first sight, all we have to do to arrive at the Newtonian theory of gravity is to take the non-relativistic limit $v/c \to 0$ in all equations. Since in the Newtonian theory the gravitational field is described by only one, time-independent, potential $\varphi(\boldsymbol{x})$, one cannot obtain the Newtonian theory by simply taking the non-relativistic limit $v/c \to 0$, since in a relativistic theory the gravitational field is described, in general, by six potentials, namely the six components $h_{ij}(t,\boldsymbol{x})$ of the metric. In this article we have used the decomposition (110). Moreover, we have shown in Section VII that only $h^i_{\parallel j}$ given by (111a) is coupled to density perturbations. By using this decomposition we have reduced the number of potentials to two, namely $\phi(t,\boldsymbol{x})$ and $\zeta(t,\boldsymbol{x})$. We have rewritten the system of equations (141) for scalar perturbations into an equivalent system (161). As a result, the perturbation to the metric, $h^i_{\parallel j}$, enters the system (161) via the trace

$$^3\!R_{(1)\parallel} = \frac{2}{c^2}\Big[2\phi^{|k}{}_{|k} - \zeta^{|k|l}{}_{|l|k} + \zeta^{|k}{}_{|k}{}^{|l}{}_{|l} + \tfrac{1}{3}\,{}^3\!R_{(0)}(3\phi + \zeta^{|k}{}_{|k})\Big], \tag{207}$$

of the spatial part of the perturbation to the Ricci tensor (114a), and via the perturbed expansion scalar (73)

$$\theta_{(1)} = \vartheta_{(1)} - \frac{1}{c^2}\big(3\dot\phi + \dot\zeta^{|k}{}_{|k}\big), \tag{208}$$

where we have used (111a). This shows explicitly that density perturbations are, in FLRW universes, described by two potentials $\phi(t,\boldsymbol{x})$ and $\zeta(t,\boldsymbol{x})$. In this section we show that in the non-relativistic limit of a *flat* FLRW universe, the potential ζ drops from the perturbation theory. For a flat (i.e. $k=0$) FLRW universe we have $^3\!R_{(0)} = 0$, as follows from (65). The perturbation to the spatial part of the Ricci scalar, (207), reduces in this case to

$$^3\!R_{(1)\parallel} = \frac{4}{c^2}\phi^{|k}{}_{|k} = -\frac{4}{c^2}\frac{\nabla^2\phi}{a^2}, \tag{209}$$

where ∇^2 is the usual Laplace operator see (51) and (151). For a flat FLRW universe, the zeroth-order equations (154) reduce to

$$\text{Constraint:} \quad 3H^2 = \kappa\varepsilon_{(0)} + \Lambda, \tag{210a}$$

$$\text{Conservation:} \quad \dot\varepsilon_{(0)} = -3H\varepsilon_{(0)}(1+w), \tag{210b}$$

$$\dot n_{(0)} = -3Hn_{(0)}. \tag{210c}$$

Upon substituting (209) into (161) and putting ${}^3R_{(0)} = 0$, we arrive at the set of perturbation equations for a flat FLRW universe:

$$\text{Constraint:} \quad H(\theta_{(1)} - \vartheta_{(1)}) + \frac{1}{c^2}\frac{\nabla^2\phi}{a^2} = \frac{4\pi G}{c^4}\left(\varepsilon^{\text{gi}}_{(1)} + \frac{\dot{\varepsilon}_{(0)}}{\dot{\theta}_{(0)}}\theta_{(1)}\right), \tag{211a}$$

$$\text{Evolution:} \quad \frac{\nabla^2\dot{\phi}}{a^2} + \frac{4\pi G}{c^2}\varepsilon_{(0)}(1+w)\vartheta_{(1)} = 0, \tag{211b}$$

$$\text{Conservation:} \quad \dot{\varepsilon}_{(1)} + 3H(\varepsilon_{(1)} + p_{(1)}) + \varepsilon_{(0)}(1+w)\theta_{(1)} = 0, \tag{211c}$$

$$\dot{\vartheta}_{(1)} + H(2 - 3\beta^2)\vartheta_{(1)} + \frac{1}{\varepsilon_{(0)}(1+w)}\frac{\nabla^2 p_{(1)}}{a^2} = 0, \tag{211d}$$

$$\dot{n}_{(1)} + 3Hn_{(1)} + n_{(0)}\theta_{(1)} = 0, \tag{211e}$$

where we have used (38), and our new definition (3a) to eliminate $\varepsilon_{(1)}$ from (161a). The scale factor of the universe $a(t)$ follows from the Einstein equations (210) via $H \equiv \dot{a}/a$. In the equations (211), the potential ζ occurs only in the quantity $\theta_{(1)}$, see (208).

We now consider the sets of equations (210) and (211) in the non-relativistic limit $v/c \to 0$. Since the spatial part $u^i_{(1)\|}$ of the fluid four-velocity is gauge dependent, (C6b), with a physical component and a non-physical gauge part, we define the non-relativistic limit $v/c \to 0$ by

$$u^i_{(1)\|\text{physical}} \equiv \frac{U^i_{(1)\|\text{physical}}}{c} \to 0, \tag{212}$$

i.e. the *physical* part of the spatial fluid velocity is negligible with respect to the speed of light. In this limit the kinetic energy per particle $\frac{1}{2}m\langle v^2\rangle = \frac{3}{2}k_{\text{B}}T \to 0$ is small compared to the rest energy mc^2 per particle, implying that the pressure $p = nk_{\text{B}}T \to 0$ ($n \neq 0$) vanishes also and that the energy density of the universe is given by $\varepsilon = nmc^2$. In other words, the non-relativistic limit (212) implies

$$p = 0, \quad \varepsilon = nmc^2, \quad w \equiv \frac{p}{\varepsilon} = 0. \tag{213}$$

In view of (213), the background equations (210) take on the simple form in the non-relativistic limit (212)

$$\text{Constraint:} \quad 3H^2 = \kappa\varepsilon_{(0)} + \Lambda, \tag{214a}$$

$$\text{Conservation:} \quad \dot{\varepsilon}_{(0)} = -3H\varepsilon_{(0)}, \tag{214b}$$

$$\dot{n}_{(0)} = -3Hn_{(0)}. \tag{214c}$$

Thus, even in the non-relativistic limit, a non-empty universe cannot be static, therefore we have $H \neq 0$. Note that in the limit (212), equations (214b) and (214c) are identical, since then $\varepsilon_{(0)} = n_{(0)}mc^2$, in view of (213).

The consequences of the limit (212) are as follows. In the limit $p \to 0$, equation (141e) reads

$$\frac{1}{c}\frac{\text{d}}{\text{d}t}\left(\varepsilon_{(0)}u^i_{(1)\|}\right) + 5H\varepsilon_{(0)}u^i_{(1)\|} = 0. \tag{215}$$

Using the background equation (214b), equation (215) takes on the simple form

$$\dot{u}^i_{(1)\|} = -2Hu^i_{(1)\|}. \tag{216}$$

The general solution of this equation is, using $H \equiv \dot{a}/a$,

$$u^i_{(1)\|\text{gauge}} = -\frac{1}{a^2(t)}\tilde{g}^{ik}\partial_k\psi(\boldsymbol{x}). \tag{217}$$

Since the physical part of $u^i_{(1)\|}$ vanishes in the limit (212), the solution (217) is a gauge mode, as follows from (C6b). Thus, in the limit (212) we are left with the non-physical quantity (217). If a quantity is proportional to the gauge function $\psi(\boldsymbol{x})$ or its partial derivatives, it may be put equal to zero without loss of physical information. If we require that in the limit (212) $u^i_{(1)\|\text{gauge}} = 0$ holds true before and after a gauge transformation we find from (217) that $\partial_k\psi(\boldsymbol{x}) = 0$ or, equivalently,

$$\psi(\boldsymbol{x}) = \psi. \tag{218}$$

In view of (218), the gauge dependent functions $\varepsilon_{(1)}$ and $n_{(1)}$ transform under a gauge transformation in the limit (212) according to (2), with constant ψ:

$$\varepsilon_{(1)} \to \hat{\varepsilon}_{(1)} = \varepsilon_{(1)} + \psi\dot{\varepsilon}_{(0)}, \tag{219a}$$

$$n_{(1)} \to \hat{n}_{(1)} = n_{(1)} + \psi\dot{n}_{(0)}. \tag{219b}$$

Since in an expanding universe the time derivatives $\dot{\varepsilon}_{(0)}(t) \neq 0$, (214b), and $\dot{n}_{(0)}(t) \neq 0$, (214c), the functions $\varepsilon_{(1)}(t, \boldsymbol{x})$ and $n_{(1)}(t, \boldsymbol{x})$ do still depend upon the gauge, so that these quantities have, also in the limit (212), no physical significance. Finally, it follows from (C2) and (218) that the gauge transformation (4) reduces in the limit (212) to the gauge transformation

$$x^0 \to x^0 - \psi, \quad x^i \to x^i - \chi^i(\boldsymbol{x}). \tag{220}$$

In other words, time and space transformations are decoupled: time coordinates may be shifted, whereas spatial coordinates may be chosen arbitrarily. Thus, in the limit (212) the general relativistic infinitesimal coordinate transformation (4), with ξ^μ given by (C2), reduces to the most general infinitesimal coordinate transformation (220) which is possible in the Newtonian theory of gravity. By now, it should be clear that the gauge problem of cosmology cannot be solved by 'fixing the gauge,' since the constant ψ and the three functions $\chi^i(\boldsymbol{x})$ cannot be determined.

We now consider the perturbation equations (211) in the limit (212). Upon substituting $\vartheta_{(1)} \equiv (u^k_{(1)\parallel})_{|k} = 0$, $w \equiv p_{(0)}/\varepsilon_{(0)} = 0$ and $p_{(1)} = 0$ into these equations we find that equation (211d) is identically satisfied, whereas the remaining equations (211) reduce to

$$\text{Constraint:} \quad \nabla^2\phi = \frac{4\pi G}{c^2}a^2\varepsilon^{\text{gi}}_{(1)}, \tag{221a}$$

$$\text{Evolution:} \quad \nabla^2\dot{\phi} = 0, \tag{221b}$$

$$\text{Conservation:} \quad \dot{\varepsilon}_{(1)} + 3H\varepsilon_{(1)} + \varepsilon_{(0)}\theta_{(1)} = 0, \tag{221c}$$

$$\dot{n}_{(1)} + 3Hn_{(1)} + n_{(0)}\theta_{(1)} = 0. \tag{221d}$$

The constraint equation (221a) can be found by subtracting $\frac{1}{6}\theta_{(1)}/\dot{H}$ times the time-derivative of the background constraint equation (214a) from the constraint equation (211a) and using that $\theta_{(0)} = 3H$, (57). This shows explicitly that (3a) is the only possible choice to construct the gauge-invariant quantity $\varepsilon^{\text{gi}}_{(1)}$. Equations (221c) and (221d) are identical in the non-relativistic limit (212) since $\varepsilon_{(1)} = n_{(1)}mc^2$, in view of (213). The quantities $\varepsilon_{(1)}$ and $n_{(1)}$, which are gauge dependent in the General Theory of Relativity, are also gauge dependent in the non-relativistic limit (212), so that equations (221c) and (221d) need not be considered: the equations (221c) and (221d) have no physical significance and can be discarded. Consequently, the perturbed expansion scalar $\theta_{(1)}$ does not occur anymore in the perturbation theory, and we are left with one potential ϕ only.

Equations (221a) and (221b) can be combined to give

$$\nabla^2\phi(\boldsymbol{x}) = \frac{4\pi G}{c^2}a^2(t)\varepsilon^{\text{gi}}_{(1)}(t, \boldsymbol{x}). \tag{222}$$

Or, equivalently,

$$\nabla^2\phi(\boldsymbol{x}) = \frac{4\pi G}{c^2}a^2(t_{\text{p}})\varepsilon^{\text{gi}}_{(1)}(t_{\text{p}}, \boldsymbol{x}), \tag{223}$$

where t_{p} indicates the present time. This Einstein equation can be rewritten in a form which is equivalent to the Poisson equation, by introducing the potential φ

$$\varphi(\boldsymbol{x}) \equiv \frac{\phi(\boldsymbol{x})}{a^2(t_{\text{p}})}. \tag{224}$$

Inserting (224) into (223) we obtain the result

$$\nabla^2\varphi(\boldsymbol{x}) = 4\pi G\frac{\varepsilon^{\text{gi}}_{(1)}(t_{\text{p}}, \boldsymbol{x})}{c^2}. \tag{225}$$

Finally, we have to check the expressions (3) or, equivalently, (196) in the limit (212). It can easily be verified that in the non-relativistic limit (212) of a flat FLRW universe, expression (196a) reduces to

$$\varepsilon^{\text{gi}}_{(1)} = -\frac{1}{2\kappa}{}^3R_{(1)\parallel}. \tag{226}$$

In view of (38) and (209), this equation is equivalent to (221a). Using (226), we find that expression (196b) reduces in the non-relativistic limit (212) to

$$n^{\rm gi}_{(1)} = n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}}\bigl(\varepsilon_{(1)} - \varepsilon^{\rm gi}_{(1)}\bigr). \tag{227}$$

Combining $\varepsilon_{(0)} = n_{(0)}mc^2$ and $\varepsilon_{(1)} = n_{(1)}mc^2$, (213), with (227) yields

$$n^{\rm gi}_{(1)} = \frac{\varepsilon^{\rm gi}_{(1)}}{mc^2}, \tag{228}$$

which is the gauge-invariant counterpart of $\varepsilon_{(1)} = n_{(1)}mc^2$. Again, we have to conclude that (3b) is the only possible choice to construct the gauge-invariant quantity $n^{\rm gi}_{(1)}$.

Combining equations (225) and (228), we arrive at the Poisson equation of the Newtonian theory of gravity:

$$\nabla^2\varphi(\boldsymbol{x}) = 4\pi G\rho_{(1)}(\boldsymbol{x}), \tag{229}$$

where

$$\rho_{(1)}(\boldsymbol{x}) \equiv \frac{\varepsilon^{\rm gi}_{(1)}(t_{\rm p},\boldsymbol{x})}{c^2}, \quad \rho_{(1)}(\boldsymbol{x}) = mn^{\rm gi}_{(1)}(t_{\rm p},\boldsymbol{x}), \tag{230}$$

is the mass density of a perturbation.

In this section we have shown that in the non-relativistic limit (212) of a flat FLRW universe, the first-order perturbation equations (161) together with the new definitions (3) reduce to the well-known Newtonian results (220), (228) and (229). Consequently, the gauge-invariant quantities $\varepsilon^{\rm gi}_{(1)}$ and $n^{\rm gi}_{(1)}$ given by (3) are indeed the energy density and particle number density perturbations.

XIII. PERTURBATION EQUATIONS FOR A FLAT FLRW UNIVERSE

In this section we derive the perturbation equations for a flat FLRW universe with a vanishing cosmological constant in its radiation-dominated, plasma-dominated and matter-dominated stages.

The six numerical values that we need in the following are: *i.* the redshift at time $t_{\rm eq}$ when the matter density has become equal to the radiation density, *ii.* the redshift at time $t_{\rm dec}$ when matter and radiation decouple, *iii.* the present value of the Hubble function, *iv.* the present value of the background radiation temperature, *v.* the age of the universe and, *vi.* the fluctuations in the background radiation at decoupling. They follow from WMAP [14–18] observations:

$$z(t_{\rm eq}) = 3176, \tag{231a}$$
$$z(t_{\rm dec}) = 1091, \tag{231b}$$
$$\mathcal{H}(t_{\rm p}) = 71.9\,\text{km/sec/Mpc} = 2.33\times10^{-18}\,\text{s}^{-1}, \tag{231c}$$
$$T_{(0)\gamma}(t_{\rm p}) = 2.725\,\text{K}, \tag{231d}$$
$$t_{\rm p} = 13.7\,\text{Gyr} = 4.32\times10^{17}\,\text{sec}, \tag{231e}$$
$$\delta_{T_\gamma}(t_{\rm dec}) \lesssim 10^{-5}, \tag{231f}$$

where we have used that $1\,\text{Mpc} = 3.0857\times10^{22}\,\text{m}$ ($1\,\text{pc} = 3.2616\,\text{ly}$).

The cosmological redshift $z(t)$ is given by

$$z(t) = \frac{a(t_{\rm p})}{a(t)} - 1, \quad a(t_{\rm p}) = 1, \tag{232}$$

where we have normalized the scale factor $a(t)$ to unity at $t = t_{\rm p}$, which is only allowed in a *flat* FLRW universe.

A. Radiation-dominated Phase

We consider the universe after the era of inflation. Let $t_{\rm rad}$ be the moment at which exponential expansion of the universe has come to an end, and the radiation-dominated era sets in. The radiation-dominated universe starts with

a temperature $T_{(0)\gamma}(t_{\text{rad}}) \approx 10^{12}\,\text{K}$ and lasts until matter-energy equality, which occurs at the redshift $z(t_{\text{eq}}) = 3176$. If we assume [see Weinberg [35], (15.5.7) for an explanation] that during the time that matter and radiation were in thermal contact the temperature of the radiation, $T_{(0)\gamma}$, dropped according to the formula $T_{(0)\gamma}(t) = A/a(t)$, where A is a constant and $a(t)$ the scale factor, we have

$$\frac{T_{(0)\gamma}(t)}{T_{(0)\gamma}(t_{\text{p}})} = \frac{a(t_{\text{p}})}{a(t)}. \tag{233}$$

This relation follows also from the Einstein equations in Section XIII A 1. We suppose that matter and radiation are still in thermal equilibrium at the end of the radiation-dominated era, i.e. at a redshift of $z(t_{\text{eq}}) = 3176$, yielding $a(t_{\text{p}})/a(t_{\text{eq}}) = 3177$. In this way we find from (231d) and (233)

$$T_{(0)\gamma}(t_{\text{eq}}) = 8657\,\text{K}. \tag{234}$$

Hence, as a rough estimate we may conclude that the universe is dominated by radiation in the temperature interval

$$10^{12}\,\text{K} \geq T_{(0)\gamma} \geq 10^{4}\,\text{K}. \tag{235}$$

If we neglect, in the radiation-dominated era, the contribution of the electrons and neutrinos, we have for the photon energy density ε

$$\varepsilon = a_{\text{B}} T_{(0)\gamma}^{4}, \tag{236}$$

where a_{B} is the black body constant

$$a_{\text{B}} = \frac{\pi^2 k_{\text{B}}^4}{15\hbar^3 c^3} = 7.5658 \times 10^{-16}\,\text{J}\,\text{m}^{-3}\,\text{K}^{-4}, \tag{237}$$

with k_{B} and $\hbar$ Boltzmann's and Planck's constant respectively. Furthermore, we could neglect the cosmological constant Λ, (160), with respect to the energy content of the universe at the matter-radiation equality time $\kappa a_{\text{B}} T_{(0)\gamma}^{4}(t_{\text{eq}}) = 1.0 \times 10^{-42}\,\text{m}^{-2}$, see equation (154a). Finally, we consider a flat universe ($k = 0$), implying, with (65), that ${}^3R_{(0)}(t) = 0$.

1. *Zeroth-order Equations*

In the radiation-dominated era, the contributions to the total energy density due to baryons, $n m_{\text{H}} c^2$, can be neglected, so that the equations of state read

$$\varepsilon(n, T) = a_{\text{B}} T_{\gamma}^{4}, \quad p(n, T) = \tfrac{1}{3} a_{\text{B}} T_{\gamma}^{4}. \tag{238}$$

Moreover, we can put $\Lambda = 0$ for reasons mentioned at the end of Section IX A. The zeroth-order equations (154) then reduce to

$$H^2 = \tfrac{1}{3}\kappa\varepsilon_{(0)}, \tag{239a}$$
$$\dot{\varepsilon}_{(0)} = -4H\varepsilon_{(0)}, \tag{239b}$$
$$\dot{n}_{(0)} = -3Hn_{(0)}. \tag{239c}$$

These equations can easily be solved, and we get the well-known results

$$H(t) = \tfrac{1}{2}\,(ct)^{-1} = H(t_{\text{rad}}) \left(\frac{t}{t_{\text{rad}}}\right)^{-1}, \tag{240a}$$
$$\varepsilon_{(0)}(t) = \frac{3}{4\kappa}\,(ct)^{-2} = \varepsilon_{(0)}(t_{\text{rad}}) \left(\frac{t}{t_{\text{rad}}}\right)^{-2}, \tag{240b}$$
$$n_{(0)}(t) = n_{(0)}(t_{\text{rad}}) \left(\frac{a(t)}{a(t_{\text{rad}})}\right)^{-3}. \tag{240c}$$

The initial values $H(t_{\text{rad}})$ and $\varepsilon_{(0)}(t_{\text{rad}})$ are related by the constraint equation (239a) taken at t_{rad}, the time at which the radiation-dominated era sets in.

Using the definition of the Hubble function $H \equiv \dot{a}/a$ we find from (240a) that

$$a(t) = a(t_{\mathrm{rad}}) \left(\frac{t}{t_{\mathrm{rad}}}\right)^{\frac{1}{2}}. \tag{241}$$

Combining expressions (236), (240b) and (241), we arrive at (233), which can be rewritten in the form

$$T_{(0)\gamma}(t) = T_{(0)\gamma}(t_{\mathrm{p}})\bigl[z(t)+1\bigr], \tag{242}$$

where we have used (232).

With (240)–(241) the coefficients b_1, b_2 and b_3, defined by (202), occurring in the equations (201) can be calculated. This will be the subject of the next section.

2. *First-order Equations*

In view of (78), we find from (238)

$$p_n = 0, \quad p_\varepsilon = \tfrac{1}{3}, \tag{243}$$

so that, according to (153), we have $w = \frac{1}{3}$ and $\beta = 1/\sqrt{3}$, see (182). The zeroth-order solutions (240) can now be substituted into the coefficients (202) of the equations (201). Since $\tilde{\boldsymbol{\nabla}} = \boldsymbol{\nabla}$ for a flat universe, the coefficients b_1, b_2 and b_3 reduce to

$$b_1 = -H, \quad b_2 = -\frac{1}{3}\frac{\nabla^2}{a^2} + \tfrac{2}{3}\kappa\varepsilon_{(0)}, \quad b_3 = 0, \tag{244}$$

where we have used (239a). For the first-order equations (201) this yields the simple forms

$$\ddot{\delta}_\varepsilon - H\dot{\delta}_\varepsilon - \left(\frac{1}{3}\frac{\nabla^2}{a^2} - \tfrac{2}{3}\kappa\varepsilon_{(0)}\right)\delta_\varepsilon = 0, \tag{245a}$$

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta_n - \tfrac{3}{4}\delta_\varepsilon\right) = 0, \tag{245b}$$

where $H \equiv \dot{a}/a$, (55). Since the right-hand side of (245a) vanishes, the evolution of density perturbations is independent of the actual value of $\delta_n - \frac{3}{4}\delta_\varepsilon$, see (201) and (203). In other words, the evolution of density perturbations is not affected by perturbations in the entropy. Equation (245b) implies that $\delta_n - \frac{3}{4}\delta_\varepsilon$ is constant in time. Consequently, during the radiation-dominated era, perturbations in the particle number density are gravitationally coupled to radiation perturbations:

$$\delta_n(t,\boldsymbol{x}) - \tfrac{3}{4}\delta_\varepsilon(t,\boldsymbol{x}) = \delta_n(t_{\mathrm{rad}},\boldsymbol{x}) - \tfrac{3}{4}\delta_\varepsilon(t_{\mathrm{rad}},\boldsymbol{x}), \tag{246}$$

where the right-hand side is constant with respect to time. From (203) we find for the perturbation of the entropy per particle:

$$s^{\mathrm{gi}}_{(1)}(t,\boldsymbol{x}) = s^{\mathrm{gi}}_{(1)}(t_{\mathrm{rad}},\boldsymbol{x}) = -\frac{a_{\mathrm{B}}T^3_{(0)\gamma}(t_{\mathrm{rad}})}{n_{(0)}(t_{\mathrm{rad}})}\Bigl[\delta_n(t_{\mathrm{rad}},\boldsymbol{x}) - \tfrac{3}{4}\delta_\varepsilon(t_{\mathrm{rad}})\Bigr], \tag{247}$$

where we have used that $T_{(0)\gamma} \propto a^{-1}$, $n_{(0)} \propto a^{-3}$ and (246). Thus, throughout the radiation-dominated era the entropy per particle is constant with respect to time. Since $n_{(0)} \neq 0$ and $n^{\mathrm{gi}}_{(1)} \neq 0$, we have $s^{\mathrm{gi}}_{(1)} \not\equiv 0$, see (194).

Equation (245a) may be solved by Fourier expansion of the function δ_ε. Writing

$$\delta_\varepsilon(t,\boldsymbol{x}) = \delta_\varepsilon(t,\boldsymbol{q})\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}}, \tag{248}$$

with $q = |\boldsymbol{q}| = 2\pi/\lambda$, where λ is the wavelength of the perturbation and $\mathrm{i}^2 = -1$, we find

$$\nabla^2\delta_\varepsilon(t,\boldsymbol{x}) = -q^2\delta_\varepsilon(t,\boldsymbol{q}), \tag{249}$$

so that the evolution equation (245a) for the amplitude $\delta_\varepsilon(t,\boldsymbol{q})$ reads

$$\ddot{\delta}_\varepsilon - H(t_{\mathrm{rad}})\left(\frac{t}{t_{\mathrm{rad}}}\right)^{-1}\dot{\delta}_\varepsilon + \left[\frac{1}{3}\frac{q^2}{a^2(t_{\mathrm{rad}})}\left(\frac{t}{t_{\mathrm{rad}}}\right)^{-1} + 2H^2(t_{\mathrm{rad}})\left(\frac{t}{t_{\mathrm{rad}}}\right)^{-2}\right]\delta_\varepsilon = 0, \tag{250}$$

where we have used (239a)–(240). This equation will be rewritten in such a way that the coefficients become dimensionless. To that end a dimensionless time variable is introduced, defined by

$$\tau \equiv \frac{t}{t_{\mathrm{rad}}}, \quad t \geq t_{\mathrm{rad}}, \tag{251}$$

with t_{rad} the time immediately after the inflationary era. This definition implies

$$\frac{\mathrm{d}^n}{c^n \mathrm{d}t^n} = \left(\frac{1}{ct_{\mathrm{rad}}}\right)^n \frac{\mathrm{d}^n}{\mathrm{d}\tau^n} = [2H(t_{\mathrm{rad}})]^n \frac{\mathrm{d}^n}{\mathrm{d}\tau^n}, \quad n = 1, 2, \dots, \tag{252}$$

where we have used (240a). Using (240a), (251) and (252), equation (250) for the density contrast $\delta_\varepsilon(\tau, \boldsymbol{q})$ can be written as

$$\delta_\varepsilon'' - \frac{1}{2\tau}\delta_\varepsilon' + \left(\frac{\mu_{\mathrm{r}}^2}{4\tau} + \frac{1}{2\tau^2}\right)\delta_\varepsilon = 0, \tag{253}$$

where a prime denotes differentiation with respect to τ. The constant μ_{r} is given by

$$\mu_{\mathrm{r}} \equiv \frac{q}{a(t_{\mathrm{rad}})} \frac{1}{H(t_{\mathrm{rad}})} \frac{1}{\sqrt{3}}. \tag{254}$$

The general solution of equation (253) is a linear combination of the functions $J_{\pm\frac{1}{2}}(\mu_{\mathrm{r}}\sqrt{\tau})\tau^{3/4}$, where $J_{+\frac{1}{2}}(x) = \sqrt{2/(\pi x)}\sin x$ and $J_{-\frac{1}{2}}(x) = \sqrt{2/(\pi x)}\cos x$ are Bessel functions of the first kind:

$$\delta_\varepsilon(\tau, \boldsymbol{q}) = \Big[A_1(\boldsymbol{q})\sin\big(\mu_{\mathrm{r}}\sqrt{\tau}\big) + A_2(\boldsymbol{q})\cos\big(\mu_{\mathrm{r}}\sqrt{\tau}\big)\Big]\sqrt{\tau}, \tag{255}$$

where the functions $A_1(\boldsymbol{q})$ and $A_2(\boldsymbol{q})$ are given by

$$A_1(\boldsymbol{q}) = \delta_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q})\sin\mu_{\mathrm{r}} - \frac{\cos\mu_{\mathrm{r}}}{\mu_{\mathrm{r}}}\left[\delta_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q}) - \frac{\dot{\delta}_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q})}{H(t_{\mathrm{rad}})}\right], \tag{256a}$$

$$A_2(\boldsymbol{q}) = \delta_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q})\cos\mu_{\mathrm{r}} + \frac{\sin\mu_{\mathrm{r}}}{\mu_{\mathrm{r}}}\left[\delta_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q}) - \frac{\dot{\delta}_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q})}{H(t_{\mathrm{rad}})}\right], \tag{256b}$$

where we have used that

$$\delta_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q}) = \delta_\varepsilon(\tau = 1, \boldsymbol{q}), \quad \dot{\delta}_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q}) = 2H(t_{\mathrm{rad}})\delta_\varepsilon'(\tau = 1, \boldsymbol{q}), \tag{257}$$

as follows from (252). The relative perturbations in the particle number density, δ_n, evolve by virtue of (246) also according to (255).

We consider the contribution of the terms of the Fourier expansion of the energy density perturbation [see (248)] in two limiting cases, namely the case of small λ (large q) and the case of large λ (small q).

For large-scale perturbations, $\lambda \to \infty$, the magnitude of the wave vector $|\boldsymbol{q}| = 2\pi/\lambda$ vanishes. Writing $\delta_\varepsilon(t) \equiv \delta_\varepsilon(t, q = 0)$ and $\dot{\delta}_\varepsilon(t) \equiv \dot{\delta}_\varepsilon(t, q = 0)$, we find from (254)–(256) that, for $t \geq t_{\mathrm{rad}}$,

$$\delta_\varepsilon(t) = -\left[\delta_\varepsilon(t_{\mathrm{rad}}) - \frac{\dot{\delta}_\varepsilon(t_{\mathrm{rad}})}{H(t_{\mathrm{rad}})}\right]\frac{t}{t_{\mathrm{rad}}} + \left[2\delta_\varepsilon(t_{\mathrm{rad}}) - \frac{\dot{\delta}_\varepsilon(t_{\mathrm{rad}})}{H(t_{\mathrm{rad}})}\right]\left(\frac{t}{t_{\mathrm{rad}}}\right)^{\frac{1}{2}}. \tag{258}$$

It is seen that the energy density contrast has two contributions to the growth rate, one proportional to t and one proportional to $t^{1/2}$. These have been found by a large number of authors. See Lifshitz and Khalatnikov [6], (8.11), Adams and Canuto [7], (4.5b), Olson [8], page 329, Peebles [9], (86.20), Kolb and Turner [10], (9.121) and Press and Vishniac [11], (33). The precise factors of proportionality, however, have not been published earlier. From the first of them we may conclude, in particular, that large-scale perturbations only grow if the initial growth rate is large enough, i.e.

$$\dot{\delta}_\varepsilon(t_{\mathrm{rad}}) \geq \delta_\varepsilon(t_{\mathrm{rad}})H(t_{\mathrm{rad}}) \quad \Rightarrow \quad \delta_\varepsilon'(\tau = 1, \boldsymbol{q}) \geq \tfrac{1}{2}\delta_\varepsilon(\tau = 1, \boldsymbol{q}), \tag{259}$$

otherwise the perturbations are decaying. For CDM and ordinary matter the same growth rate $\delta_n \propto t$ is found in the literature for super-horizon perturbations, see, for example, the textbook of Padmanabhan [39], Section 4.4. Thus, for large-scale perturbations, our treatise corroborates the outcomes found in the literature on the subject.

We now come to the second case. In the small-scale limit $\lambda \to 0$ (or, equivalently, $|\boldsymbol{q}| \to \infty$) we find, using (254)–(256), that

$$\delta_\varepsilon(t, \boldsymbol{q}) \approx \delta_\varepsilon(t_{\mathrm{rad}}, \boldsymbol{q}) \left(\frac{t}{t_{\mathrm{rad}}}\right)^{\frac{1}{2}} \cos\left[\mu_{\mathrm{r}} - \mu_{\mathrm{r}}\left(\frac{t}{t_{\mathrm{rad}}}\right)^{\frac{1}{2}}\right]. \tag{260}$$

We see that in the limit of small λ, the contribution to the growth rate is smaller than the leading term in the expression (258). Physically, this can be understood: on small scales, the pressure gradients $|\boldsymbol{\nabla} p| \approx p/\lambda$ are much higher than on large scales.

In Section XVI we review the standard results. Comparing our result (260) with the standard result (352) found in the literature, we observe that we obtain oscillating solutions with an increasing amplitude, whereas the standard equation (350) yields oscillating solutions (352) with a decreasing amplitude.

The above calculated behavior of density perturbations in the radiation-dominated universe is important for star formation in the era after decoupling of matter and radiation: the oscillating growth shows up in the cosmic background radiation as random fluctuations on different scales and amplitudes (i.e. intensities in the background radiation). As we will show in Section XV, density perturbations yield massive stars.

B. Plasma Era

The so-called plasma era sets in at time t_{eq}, when the energy density of ordinary matter equals the energy density of radiation, i.e, when $n_{(0)}(t_{\mathrm{eq}})mc^2 = a_{\mathrm{B}}T^4_{(0)\gamma}(t_{\mathrm{eq}})$, and ends at time t_{dec}, the time of decoupling of matter and radiation. In the plasma era the matter-radiation mixture can be characterized by the equations of state, see Kodama and Sasaki [25] Chapter V,

$$\varepsilon(n, T) = nmc^2 + a_{\mathrm{B}}T^4_\gamma, \quad p(n, T) = \tfrac{1}{3}a_{\mathrm{B}}T^4_\gamma, \tag{261}$$

where we have not taken into account the contributions to the pressure from ordinary matter, $m = m_{\mathrm{H}}$, or CDM, $m = m_{\mathrm{cdm}}$, since these contributions are vanishingly small in comparison to the radiation energy density. Using (A5), we find

$$p_n = -\tfrac{1}{3}mc^2, \quad p_\varepsilon = \tfrac{1}{3}. \tag{262}$$

Substituting the expressions (261) and (262) into the entropy equation (201b) and integrating the resulting equation yields

$$\delta_n(t, \boldsymbol{x}) - \frac{\delta_\varepsilon(t, \boldsymbol{x})}{1 + w(t)} = \left[\delta_n(t_{\mathrm{eq}}, \boldsymbol{x}) - \frac{\delta_\varepsilon(t_{\mathrm{eq}}, \boldsymbol{x})}{1 + w(t_{\mathrm{eq}})}\right] \exp\left[-\int_{ct_{\mathrm{eq}}}^{ct} \frac{H(\tau)n_{(0)}(\tau)mc^2}{n_{(0)}(\tau)mc^2 + \frac{4}{3}a_{\mathrm{B}}T^4_{(0)\gamma}(\tau)}\mathrm{d}\tau\right]. \tag{263}$$

It is well-known that ordinary matter perturbations are coupled to perturbations in the radiation density. This coupling is attributed to the high radiation pressure in the radiation-dominated and plasma eras. In addition to this coupling, expression (263) shows that perturbations in the particle number density, δ_n, are also *gravitationally* coupled to perturbations in the total energy density, δ_ε. This has considerable consequences for the growth of perturbations in CDM. Since the energy density of CDM can be written as $n_{(0)}m_{\mathrm{cdm}}c^2$, (263) implies that also density perturbations in CDM are gravitationally coupled to perturbations in the radiation energy density. Consequently, CDM perturbations can, just as perturbations in ordinary matter, start to grow only after decoupling. In other words, CDM and ordinary matter behave gravitationally in exactly the same way. This may rule out CDM as a means to facilitate the formation of structure in the universe. The same conclusion, on different grounds, has also been reached by Nieuwenhuizen *et al.* [5].

C. Era after Decoupling of Matter and Radiation

In the era after decoupling of matter and radiation, we have to distinguish between the matter temperature and the radiation temperature. The radiation temperature $T_{(0)\gamma}$ evolves as (233), whereas the matter temperature $T_{(0)}$ evolves according to (277). Once protons and electrons recombine to yield hydrogen, the radiation pressure becomes negligible, and the equations of state reduce to those of a non-relativistic monatomic perfect gas [Weinberg [35], equations (15.8.20) and (15.8.21)]

$$\varepsilon(n, T) = nm_{\mathrm{H}}c^2 + \tfrac{3}{2}nk_{\mathrm{B}}T, \quad p(n, T) = nk_{\mathrm{B}}T, \tag{264}$$

where $k_{\mathrm{B}} = 1.3806504 \times 10^{-23}\,\mathrm{J\,K^{-1}}$ is Boltzmann's constant, m_{H} the mass of a proton, and T the temperature of the *matter*. Since the energy density in (264) is not of the form $\varepsilon = \varepsilon(n)$, see Section X D, density perturbations cannot be adiabatic.

1. Zeroth-order Equations

The maximum gas temperature occurs around time t_{dec} of the decoupling of matter and radiation and is equal to the radiation temperature $T_{(0)\gamma}(t_{\mathrm{dec}})$. Using (231)–(233), we get for the temperature at the time of decoupling

$$T_{(0)}(t_{\mathrm{dec}}) = T_{(0)\gamma}(t_{\mathrm{dec}}) = 2976\,\mathrm{K}. \tag{265}$$

Since the universe cools down during its expansion, it follows from (264) that

$$\frac{p}{\varepsilon} \approx \frac{k_{\mathrm{B}} T_{(0)}(t)}{m_{\mathrm{H}} c^2} \leq \frac{k_{\mathrm{B}} T_{(0)}(t_{\mathrm{dec}})}{m_{\mathrm{H}} c^2} = 2.73 \times 10^{-10}, \quad t \geq t_{\mathrm{dec}}. \tag{266}$$

Hence, the pressure is negligible with respect to the energy density. This implies that, to a good approximation, $\varepsilon_{(0)} \pm p_{(0)} \approx \varepsilon_{(0)}$ and $\varepsilon_{(0)} \approx n_{(0)} m_{\mathrm{H}} c^2$. Hence, in an unperturbed flat FLRW universe the pressure can, in the background equations, be neglected with respect to the energy density. The above facts yield that the Einstein equations and conservation laws (154) for a flat FLRW universe reduce to

$$H^2 = \tfrac{1}{3}\kappa\varepsilon_{(0)}, \tag{267a}$$
$$\dot{\varepsilon}_{(0)} = -3H\varepsilon_{(0)}, \tag{267b}$$
$$\dot{n}_{(0)} = -3Hn_{(0)}, \tag{267c}$$

where we have put the cosmological constant Λ equal to zero. The solutions of (267) are

$$H(t) = \tfrac{2}{3}(ct)^{-1} = H(t_{\mathrm{mat}}) \left(\frac{t}{t_{\mathrm{mat}}}\right)^{-1}, \tag{268a}$$
$$\varepsilon_{(0)}(t) = \frac{4}{3\kappa}(ct)^{-2} = \varepsilon_{(0)}(t_{\mathrm{mat}}) \left(\frac{t}{t_{\mathrm{mat}}}\right)^{-2}, \tag{268b}$$
$$n_{(0)}(t) = n_{(0)}(t_{\mathrm{mat}}) \left(\frac{a(t)}{a(t_{\mathrm{mat}})}\right)^{-3}. \tag{268c}$$

The scale factor $a(t)$ after decoupling, however, has a time dependence which differs from that of the radiation-dominated era (241). From (268a) we find

$$a(t) = a(t_{\mathrm{mat}}) \left(\frac{t}{t_{\mathrm{mat}}}\right)^{\frac{2}{3}}, \tag{269}$$

where t_{mat} is some initial time after decoupling of matter and radiation:

$$t_{\mathrm{dec}} \leq t_{\mathrm{mat}} \leq t_{\mathrm{p}}. \tag{270}$$

The initial values $H(t_{\mathrm{mat}})$ and $\varepsilon_{(0)}(t_{\mathrm{mat}})$ are related by the constraint equation (267a) taken at $t = t_{\mathrm{mat}}$. With (268)–(269) the coefficients (202) of the perturbation equations (201) are known functions of time.

2. First-order Equations

We first remark that, in the study of the evolution of density perturbations, we may not neglect the pressure with respect to the energy density. The case of a pressureless perfect fluid is already thoroughly discussed in Section XII on the non-relativistic limit. We neglect $k_{\mathrm{B}} T_{(0)}/(m_{\mathrm{H}} c^2)$ with respect to terms of the order unity in the final expressions.

Using equations (A5) we find from the equations of state (264)

$$p_{\varepsilon} = \tfrac{2}{3}, \quad p_n = -\tfrac{2}{3} m_{\mathrm{H}} c^2, \tag{271}$$

From (182) it follows that

$$\beta(t) = \sqrt{\frac{2}{3}\left[1 - \frac{m_{\mathrm{H}}c^2}{m_{\mathrm{H}}c^2 + \frac{5}{2}k_{\mathrm{B}}T_{(0)}(t)}\right]}. \tag{272}$$

Since $k_{\mathrm{B}}T_{(0)}(t) \ll m_{\mathrm{H}}c^2$, (266), we find, to a good approximation

$$\beta(t) \approx \frac{v_{\mathrm{s}}(t)}{c} = \sqrt{\frac{5}{3}\frac{k_{\mathrm{B}}T_{(0)}(t)}{m_{\mathrm{H}}c^2}}, \tag{273}$$

where v_{s} is the speed of sound. Differentiating (273) with respect to time yields

$$\frac{\dot{\beta}}{\beta} = \frac{\dot{T}_{(0)}}{2T_{(0)}}. \tag{274}$$

For the time development of the matter temperature it is found from (192) and (264) that

$$\dot{T}_{(0)} = -2HT_{(0)}. \tag{275}$$

Combining (274) and (275) results in

$$\frac{\dot{\beta}}{\beta} = -H. \tag{276}$$

For the evolution of the matter temperature it is found from (275) that

$$T_{(0)}(t) = T_{(0)}(t_{\mathrm{mat}})\left(\frac{a(t)}{a(t_{\mathrm{mat}})}\right)^{-2} = T_{(0)}(t_{\mathrm{mat}})\left(\frac{z(t)+1}{z(t_{\mathrm{mat}})+1}\right)^{2}, \tag{277}$$

where we have used that $H \equiv \dot{a}/a$ and (232).

We now consider equation (201b). Using (271) it is found for this equation

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta_n - \delta_\varepsilon\right) = -2H\left(\delta_n - \delta_\varepsilon\right). \tag{278}$$

The general solution of equation (278) is, using also (269),

$$\Delta(t,\boldsymbol{x}) = \Delta(t_{\mathrm{mat}},\boldsymbol{x})\left(\frac{a(t)}{a(t_{\mathrm{mat}})}\right)^{-2} = \Delta(t_{\mathrm{mat}},\boldsymbol{x})\left(\frac{t}{t_{\mathrm{mat}}}\right)^{-\frac{4}{3}}, \tag{279}$$

where the quantity $\Delta(t,\boldsymbol{x})$ is defined by

$$\Delta(t,\boldsymbol{x}) \equiv \delta_n(t,\boldsymbol{x}) - \delta_\varepsilon(t,\boldsymbol{x}). \tag{280}$$

The combined First and Second Laws of thermodynamics (203) reads for an equation of state (264)

$$T_{(0)}s^{\mathrm{gi}}_{(1)} = -(m_{\mathrm{H}}c^2 + \tfrac{3}{2}k_{\mathrm{B}}T_{(0)})\Delta - k_{\mathrm{B}}T_{(0)}\delta_n. \tag{281}$$

Using (264) we find from (206) and (271) that the relative pressure perturbation is given by

$$\delta_p = \delta_n - \tfrac{2}{3}\Delta\frac{m_{\mathrm{H}}c^2 + \frac{3}{2}k_{\mathrm{B}}T_{(0)}}{k_{\mathrm{B}}T_{(0)}}. \tag{282}$$

The quantity $\Delta(t,\boldsymbol{x})$ can now be rewritten as

$$\Delta = -\frac{\frac{3}{2}k_{\mathrm{B}}T_{(0)}}{m_{\mathrm{H}}c^2 + \frac{3}{2}k_{\mathrm{B}}T_{(0)}}(\delta_p - \delta_n). \tag{283}$$

Eliminating $\Delta(t,\boldsymbol{x})$ from (281) with the help of (283) we find for the heat transfer to a perturbation from its surroundings

$$T_{(0)}s^{\mathrm{gi}}_{(1)} = \tfrac{1}{2}k_{\mathrm{B}}T_{(0)}\left(3\delta_p - 5\delta_n\right). \tag{284}$$

A density perturbation loses some of its heat energy to its surroundings if $T_{(0)}s^{\mathrm{gi}}_{(1)} < 0$, i.e. if

$$\delta_n(t,\boldsymbol{x}) > \tfrac{3}{5}\delta_p(t,\boldsymbol{x}). \tag{285}$$

If this condition is fulfilled, then a density perturbation may grow provided that its internal gravity is strong enough. The relative perturbation in the *matter* temperature and its time-derivative follow from (205), (264), (275), (278) and (280). We get,

$$\delta_T = -\tfrac{2}{3}\Delta\frac{m_{\mathrm{H}}c^2 + \frac{3}{2}k_{\mathrm{B}}T_{(0)}}{k_{\mathrm{B}}T_{(0)}}, \quad \dot{\delta}_T = 2H\Delta, \tag{286}$$

so that for $\delta_n > \delta_\varepsilon$ we have $\delta_T < 0$, and the matter temperature perturbation increases slowly, $\dot{\delta}_T \gtrsim 0$. Combining (282) and (286), we arrive at the simple relation

$$\delta_p = \delta_T + \delta_n, \tag{287}$$

i.e. both temperature fluctuations and perturbations in the particle number density contribute to perturbations in the pressure. Since $m_{\mathrm{H}}c^2 \gg k_{\mathrm{B}}T_{(0)}$ during the era after decoupling, it follows from (277), (279) and (286) that we may assume that

$$\delta_T(t,\boldsymbol{x}) \approx \delta_T(t_{\mathrm{mat}},\boldsymbol{x}), \quad \dot{\delta}_T(t,\boldsymbol{x}) \approx 0, \tag{288}$$

to a very good approximation. Apparently the heat exchange is so efficient that the gas remains at a nearly constant temperature during the linear phase of the gravitational collapse. With (287) this implies that

$$\delta_p(t,\boldsymbol{x}) \approx \delta_T(t_{\mathrm{mat}},\boldsymbol{x}) + \delta_n(t,\boldsymbol{x}). \tag{289}$$

Thus, for a growing density perturbation the internal pressure will be increased.

Expression (286) can be approximated by

$$\delta_n(t,\boldsymbol{x}) - \delta_\varepsilon(t,\boldsymbol{x}) \equiv \Delta(t,\boldsymbol{x}) \approx -\frac{3}{2}\frac{k_{\mathrm{B}}T_{(0)}(t)}{m_{\mathrm{H}}c^2}\delta_T(t_{\mathrm{mat}},\boldsymbol{x}) = -\frac{9}{10}\frac{v_{\mathrm{s}}^2(t)}{c^2}\delta_T(t_{\mathrm{mat}},\boldsymbol{x}), \tag{290}$$

where we have used that $m_{\mathrm{H}}c^2 \gg k_{\mathrm{B}}T_{(0)}$. Note that (290) is a solution of (278), since $T_{(0)}(t) \propto a^{-2}(t)$, (277).

Finally, using (289) the heat transfer (284) can now be rewritten as

$$T_{(0)}(t)s^{\mathrm{gi}}_{(1)}(t,\boldsymbol{x}) \approx \tfrac{1}{2}k_{\mathrm{B}}T_{(0)}(t)\big[3\delta_T(t_{\mathrm{mat}},\boldsymbol{x}) - 2\delta_n(t,\boldsymbol{x})\big]. \tag{291}$$

This concludes the discussion of equation (201b).

We now consider (201a). After substituting (271), (273) and (276) into the coefficients (202) it is found that

$$b_1 = 3H, \quad b_2 = -\tfrac{5}{6}\kappa\varepsilon_{(0)} - \frac{v_{\mathrm{s}}^2}{c^2}\frac{\nabla^2}{a^2}, \quad b_3 = -\frac{2}{3}\frac{\nabla^2}{a^2}, \tag{292}$$

considering that for a flat universe $\tilde{\nabla}^2 = \nabla^2$. In the derivation of b_3 we have used that $\beta^2 \ll 1$, as follows from (266) and (273). For the evolution equation for density perturbations, (201a), this results in the simple form

$$\ddot{\delta}_\varepsilon + 3H\dot{\delta}_\varepsilon - \left(\frac{v_{\mathrm{s}}^2}{c^2}\frac{\nabla^2}{a^2} + \tfrac{5}{6}\kappa\varepsilon_{(0)}\right)\delta_\varepsilon = \frac{3}{5}\frac{v_{\mathrm{s}}^2}{c^2}\frac{\nabla^2}{a^2}\delta_T(t_{\mathrm{mat}},\boldsymbol{x}), \tag{293}$$

where we have used the result (290). Using (248) and (249) the evolution equation for the amplitude $\delta_\varepsilon(t,\boldsymbol{q})$ can be rewritten as

$$\ddot{\delta}_\varepsilon + 3H(t_{\mathrm{mat}})\left(\frac{t}{t_{\mathrm{mat}}}\right)^{-1}\dot{\delta}_\varepsilon + H^2(t_{\mathrm{mat}})\left[\mu_{\mathrm{m}}^2\left(\frac{t}{t_{\mathrm{mat}}}\right)^{-\frac{8}{3}} - \frac{5}{2}\left(\frac{t}{t_{\mathrm{mat}}}\right)^{-2}\right]\delta_\varepsilon = -\frac{3}{5}H^2(t_{\mathrm{mat}})\mu_{\mathrm{m}}^2\left(\frac{t}{t_{\mathrm{mat}}}\right)^{-\frac{8}{3}}\delta_T(t_{\mathrm{mat}},\boldsymbol{q}), \tag{294}$$

where we have incorporated (267)–(269), (273) and (277). The constant μ_m is given by

$$\mu_\mathrm{m} \equiv \frac{q}{a(t_\mathrm{mat})}\frac{1}{H(t_\mathrm{mat})}\frac{v_\mathrm{s}(t_\mathrm{mat})}{c}, \quad v_\mathrm{s}(t_\mathrm{mat}) = \sqrt{\frac{5}{3}\frac{k_\mathrm{B}T_{(0)}(t_\mathrm{mat})}{m_\mathrm{H}}}. \tag{295}$$

Using the dimensionless time variable

$$\tau \equiv \frac{t}{t_\mathrm{mat}}, \quad t_\mathrm{mat} \le t \le t_\mathrm{p}, \tag{296}$$

it is found from (268a) that

$$\frac{\mathrm{d}^n}{c^n \mathrm{d}t^n} = \left(\frac{1}{ct_\mathrm{mat}}\right)^n \frac{\mathrm{d}^n}{\mathrm{d}\tau^n} = \left[\frac{3}{2}H(t_\mathrm{mat})\right]^n \frac{\mathrm{d}^n}{\mathrm{d}\tau^n}, \quad n = 1, 2, \ldots . \tag{297}$$

Using this expression, equation (294) can be rewritten in the form

$$\delta_\varepsilon'' + \frac{2}{\tau}\delta_\varepsilon' + \left(\frac{4}{9}\frac{\mu_\mathrm{m}^2}{\tau^{8/3}} - \frac{10}{9\tau^2}\right)\delta_\varepsilon = -\frac{4}{15}\frac{\mu_\mathrm{m}^2}{\tau^{8/3}}\delta_T(t_\mathrm{mat}, \boldsymbol{q}), \tag{298}$$

where we have also used (290). In equation (298), a prime denotes differentiation with respect to τ.

It is of convenience for the numerical integration of equation (298) to express the dimensionless time variable τ, (296), in the cosmological redshift $z(t)$. Using (232) and (269), we get

$$\tau = \left(\frac{a(t)}{a(t_\mathrm{mat})}\right)^{\frac{3}{2}} = \left(\frac{z(t_\mathrm{mat}) + 1}{z(t) + 1}\right)^{\frac{3}{2}}. \tag{299}$$

The integration will then be halted if either the time variable τ has reached the value τ_end for which $z = 0$, i.e.

$$\tau_\mathrm{end} = \left[z(t_\mathrm{mat}) + 1\right]^{3/2}, \tag{300}$$

or when $|\delta_\varepsilon(\tau, \boldsymbol{q})| = 1$ has been reached for $\tau < \tau_\mathrm{end}$. In Section XV, we solve numerically the inhomogeneous equation (298). First, however, we consider the homogeneous part of the evolution equation (298).

3. *General Solution of the Evolution Equation*

In this section we calculate the exact solution of equation (298). To that end, we replace the independent variable τ in this equation by the new independent variable

$$x(\tau) \equiv 2\mu_\mathrm{m}\tau^{-1/3}, \tag{301}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = -\tfrac{2}{3}\mu_\mathrm{m}\tau^{-4/3}\frac{\mathrm{d}}{\mathrm{d}x}, \qquad \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} = \tfrac{8}{9}\mu_\mathrm{m}\tau^{-7/3}\frac{\mathrm{d}}{\mathrm{d}x} + \tfrac{4}{9}\mu_\mathrm{m}^2\tau^{-8/3}\frac{\mathrm{d}^2}{\mathrm{d}x^2}. \tag{302}$$

Using the variable (301), equation (298) can be rewritten in a much more tractable form

$$\frac{\mathrm{d}^2\delta_\varepsilon}{\mathrm{d}x^2} - \frac{2}{x}\frac{\mathrm{d}\delta_\varepsilon}{\mathrm{d}x} + \left(1 - \frac{10}{x^2}\right)\delta_\varepsilon = -\tfrac{3}{5}\delta_T(t_\mathrm{mat}, \boldsymbol{q}). \tag{303}$$

The general solution of this equation is

$$\delta_\varepsilon(x) = \left[A_1 J_{+\frac{7}{2}}(x) + A_2 J_{-\frac{7}{2}}(x)\right]x^{3/2} - \frac{3}{5}\left(1 + \frac{10}{x^2}\right)\delta_T(t_\mathrm{mat}, \boldsymbol{q}), \tag{304}$$

where A_1 and A_2 are the constants of integration and $J_{\pm\nu}(x)$ are the Bessel functions of the first kind:

$$J_{+\frac{7}{2}}(x) = \sqrt{\frac{2}{\pi}}\left[(x^3 - 15x)\cos x - (6x^2 - 15)\sin x\right]x^{-7/2}, \tag{305a}$$

$$J_{-\frac{7}{2}}(x) = \sqrt{\frac{2}{\pi}}\left[(x^3 - 15x)\sin x + (6x^2 - 15)\cos x\right]x^{-7/2}, \tag{305b}$$

with the asymptotic expressions for $x \to 0$

$$J_{+\frac{7}{2}}(x) \approx \sqrt{\frac{2}{\pi}} \frac{x^{7/2}}{105}, \qquad J_{-\frac{7}{2}}(x) \approx -\sqrt{\frac{2}{\pi}} \frac{15}{x^{7/2}}. \tag{306}$$

Transforming back from x to τ, we arrive at the general solution of equation (298)

$$\delta_\varepsilon(\tau, \boldsymbol{q}) = \left[B_1(\boldsymbol{q}) J_{+\frac{7}{2}}\left(2\mu_{\mathrm{m}}\tau^{-1/3}\right) + B_2(\boldsymbol{q}) J_{-\frac{7}{2}}\left(2\mu_{\mathrm{m}}\tau^{-1/3}\right)\right] \tau^{-1/2} - \frac{3}{5}\left(1 + \frac{5\tau^{2/3}}{2\mu_{\mathrm{m}}^2}\right) \delta_T(t_{\mathrm{mat}}, \boldsymbol{q}), \tag{307}$$

where $B_1(\boldsymbol{q})$ and $B_2(\boldsymbol{q})$ are arbitrary functions. The first two terms are the solution of the homogeneous equation and the last term is the particular solution. The solution of the homogeneous equation follows from $\delta_T(t_{\mathrm{mat}}, \boldsymbol{q}) = 0$ and has the property

$$\delta_T(t, \boldsymbol{x}) = 0 \quad \Leftrightarrow \quad \delta_n(t, \boldsymbol{x}) = \delta_\varepsilon(t, \boldsymbol{x}) = \delta_p(t, \boldsymbol{x}), \tag{308}$$

where we have used (280), (283) and (286). Thus, if $\delta_\varepsilon = \delta_n$ then pressure perturbations counteract the growth or decay of a density perturbation in such a way that $\delta_n = \delta_p$ and temperature perturbations in the matter do not occur. This is always the case in the standard Newtonian perturbation theory given by equation (357), since in this theory one has $\delta_n = \delta_\varepsilon$. In the literature about linear cosmological perturbations, equation (278) does not exist, and the standard second-order evolution equation (357) is homogeneous. Since $m_{\mathrm{H}}c^2 \gg k_{\mathrm{B}} T_{(0)}$, one is forced to take $\delta_n = \delta_\varepsilon$. It follows from (284) and (308) that, in this case, the heat transfer to a perturbation is given by

$$T_{(0)} s_{(1)}^{\mathrm{gi}} = -k_{\mathrm{B}} T_{(0)} \delta_n. \tag{309}$$

This implies that a density perturbation with $\delta_n > 0$ loses some of its internal heat energy to its surroundings, so that it may grow. The heat transfer $T_{(0)} s_{(1)}^{\mathrm{gi}}$ from the perturbation to its surroundings is very small. In contrast to the standard theory, our treatise allows $\delta_n \neq \delta_\varepsilon$, which may result in a larger heat loss and, hence, a faster growth rate. This will be studied in Section XV.

The evolution of density perturbations can be studied by imposing initial conditions $\delta_\varepsilon(t_{\mathrm{mat}}, \boldsymbol{q})$ and $\dot{\delta}_\varepsilon(t_{\mathrm{mat}}, \boldsymbol{q})$ on the general solution (307). Since the resulting expression is far too complicated, we investigate the evolution of density perturbations by solving equation (298) numerically in Section XV. In this section we only consider the two limiting cases of large-scale and small-scale perturbations.

In the large-scale limit $|\boldsymbol{q}| \to 0$, $\lambda \to \infty$ (i.e. larger than the horizon, see Appendix D), it is found that, transforming back from τ to t,

$$\delta_\varepsilon(t) = \frac{1}{7}\left[5\delta_\varepsilon(t_{\mathrm{mat}}) + \frac{2\dot{\delta}_\varepsilon(t_{\mathrm{mat}})}{H(t_{\mathrm{mat}})}\right]\left(\frac{t}{t_{\mathrm{mat}}}\right)^{\frac{2}{3}} + \frac{2}{7}\left[\delta_\varepsilon(t_{\mathrm{mat}}) - \frac{\dot{\delta}_\varepsilon(t_{\mathrm{mat}})}{H(t_{\mathrm{mat}})}\right]\left(\frac{t}{t_{\mathrm{mat}}}\right)^{-\frac{5}{3}}. \tag{310}$$

Thus, for large-scale perturbations, the initial value $\delta_T(t_{\mathrm{mat}}, \boldsymbol{q})$ does not play a role during the evolution: large-scale perturbations evolve only under the influence of gravity. The solution proportional to $t^{2/3}$ is a standard result. Since δ_ε is gauge-invariant, the standard non-physical gauge mode proportional to t^{-1} is absent from our theory. Instead, a physical mode proportional to $t^{-5/3}$ is found. This mode has also been found by Bardeen [12], Table I, and by Mukhanov *et al.* [13], expression (5.33). In order to arrive at the $t^{-5/3}$ mode, Bardeen has to use the 'uniform Hubble constant gauge.' In our treatise, however, the Hubble function is automatically uniform, without any additional gauge condition, see (25).

In the small-scale limit $\lambda \to 0$ or, equivalently, $|\boldsymbol{q}| \to \infty$, we find, transforming back from τ to t,

$$\delta_\varepsilon(t, \boldsymbol{q}) \approx -\tfrac{3}{5}\delta_T(t_{\mathrm{mat}}, \boldsymbol{q}) + \left(\frac{t}{t_{\mathrm{mat}}}\right)^{-1/3}\left[\tfrac{3}{5}\delta_T(t_{\mathrm{mat}}, \boldsymbol{q}) + \delta_\varepsilon(t_{\mathrm{mat}}, \boldsymbol{q})\right] \cos\left[2\mu_{\mathrm{m}} - 2\mu_{\mathrm{m}}\left(\frac{t}{t_{\mathrm{mat}}}\right)^{-1/3}\right]. \tag{311}$$

Thus, small-scale density perturbations oscillate with a decaying amplitude which is smaller than unity so that the non-linear regime will never be reached.

For scales between the two extremes discussed above the growth of the perturbations can be considerable as we will show in Section XV.

XIV. STAR FORMATION: BASIC EQUATIONS

This section is a preparation for the numerical solution of the evolution equation (298). We express the constant μ_m and the star mass $M(t_\mathrm{mat})$ in the observable quantities $z(t_\mathrm{dec})$, $\mathcal{H}(t_\mathrm{p})$ and $T_{(0)}(t_\mathrm{dec})$, the initial redshift $z(t_\mathrm{mat})$, where t_mat is some initial time (270) at which a density perturbation starts to contract, and the physical dimensions λ_mat of a density perturbation. Finally, we study the influence of the particle mass m and the initial time t_mat on the mass of a star.

A. Initial Values and the Mass of a Star

Writing $q = 2\pi/\lambda$, where $\lambda a(t_\mathrm{mat})$ is the physical scale of a perturbation at time t_mat and $\lambda a(t_\mathrm{p}) \equiv \lambda$ is the physical scale as measured at the present time t_p, we get for (295)

$$\mu_\mathrm{m} = \frac{2\pi}{\lambda_\mathrm{mat}} \frac{1}{H(t_\mathrm{mat})} \sqrt{\frac{5}{3} \frac{k_\mathrm{B} T_{(0)}(t_\mathrm{mat})}{m_\mathrm{H} c^2}}, \quad \lambda_\mathrm{mat} \equiv \lambda a(t_\mathrm{mat}). \tag{312}$$

The Hubble function $H(t_\mathrm{mat})$ follows from (232), (268a) and (269). We find

$$H(t_\mathrm{mat}) = H(t_\mathrm{p}) \bigl[z(t_\mathrm{mat}) + 1\bigr]^{3/2}. \tag{313}$$

From (277), we get

$$T_{(0)}(t_\mathrm{mat}) = T_{(0)}(t_\mathrm{dec}) \left(\frac{z(t_\mathrm{mat}) + 1}{z(t_\mathrm{dec}) + 1} \right)^2. \tag{314}$$

Using (313) and (314), expression (312) can be rewritten as

$$\mu_\mathrm{m} = \frac{2\pi}{\lambda_\mathrm{mat}} \frac{1}{\mathcal{H}(t_\mathrm{p}) \bigl[z(t_\mathrm{dec}) + 1\bigr] \sqrt{z(t_\mathrm{mat}) + 1}} \sqrt{\frac{5}{3} \frac{k_\mathrm{B} T_{(0)}(t_\mathrm{dec})}{m_\mathrm{H}}}, \tag{315}$$

where we have used (54). With (315) we have expressed μ_m in the observable quantities $\mathcal{H}(t_\mathrm{p})$, $z(t_\mathrm{dec})$ and $T_{(0)}(t_\mathrm{dec})$. The matter temperature just after decoupling, $T_{(0)}(t_\mathrm{dec})$, is given by (265). Using also (231) we get

$$\mu_\mathrm{m} = \frac{512.0}{\lambda_\mathrm{mat} \sqrt{z(t_\mathrm{mat}) + 1}}, \quad \lambda_\mathrm{mat} \text{ in pc}, \tag{316}$$

where we have used that $1\,\mathrm{pc} = 3.0857 \times 10^{16}\,\mathrm{m}$ ($1\,\mathrm{pc} = 3.2616\,\mathrm{ly}$).

The mass $M(t_\mathrm{mat})$ of a spherical density perturbation with radius $\frac{1}{2}\lambda_\mathrm{mat}$ is given by

$$M(t_\mathrm{mat}) = \frac{4\pi}{3} \left(\tfrac{1}{2}\lambda_\mathrm{mat}\right)^3 n_{(0)}(t_\mathrm{mat}) m_\mathrm{H}, \tag{317}$$

where m_H is the proton mass. The particle number density can be calculated from the value at the end of the radiation-dominated era. By definition, at the end of the radiation-domination era the matter energy density $n_{(0)} m_\mathrm{H} c^2$ equals the energy density of the radiation. Hence

$$n_{(0)}(t_\mathrm{eq}) m_\mathrm{H} c^2 = a_\mathrm{B} T^4_{(0)\gamma}(t_\mathrm{eq}). \tag{318}$$

Since $n_{(0)} \propto a^{-3}$ and $T_{(0)\gamma} \propto a^{-1}$, we find, using (232), the particle number density at time t_mat

$$n_{(0)}(t_\mathrm{mat}) = \frac{a_\mathrm{B} T^4_{(0)\gamma}(t_\mathrm{p})}{m_\mathrm{H} c^2} \bigl[z(t_\mathrm{eq}) + 1\bigr] \bigl[z(t_\mathrm{mat}) + 1\bigr]^3. \tag{319}$$

Combining (317) and (319), we get

$$M(t_\mathrm{mat}) = \frac{4\pi}{3} \left(\tfrac{1}{2}\lambda_\mathrm{mat}\right)^3 \frac{a_\mathrm{B} T^4_{(0)\gamma}(t_\mathrm{p})}{c^2} \bigl[z(t_\mathrm{eq}) + 1\bigr] \bigl[z(t_\mathrm{mat}) + 1\bigr]^3. \tag{320}$$

Using that one solar mass is 1.98892×10^{30} kg, we find from (231)

$$M(t_{\mathrm{mat}}) = 1.141 \times 10^{-8} \lambda_{\mathrm{mat}}^3 \bigl[z(t_{\mathrm{mat}}) + 1\bigr]^3 \mathrm{M}_\odot, \quad \lambda_{\mathrm{mat}} \text{ in pc.} \tag{321}$$

This expression will be used to convert the scale λ_{mat} of a perturbation, which starts to contract at a redshift of $z(t_{\mathrm{mat}})$, into its mass. The latter is expressed in units of the solar mass. It should be stressed here that the numeric factors in expressions (316) and (321) hold true only for a fluid consisting of protons. In Section XIV B we investigate the influence of the particle mass on the mass of a star.

In order to solve equation (298), we need the initial values $\delta_\varepsilon(t_{\mathrm{mat}}, \boldsymbol{q})$, $\delta_\varepsilon'(t_{\mathrm{mat}}, \boldsymbol{q})$ and $\delta_T(t_{\mathrm{mat}}, \boldsymbol{q})$. From (290) it follows that, since $|\delta_T(t, \boldsymbol{q})| \leq 1$, we must have

$$|\delta_n(t, \boldsymbol{q}) - \delta_\varepsilon(t, \boldsymbol{q})| \approx \tfrac{3}{2}|\delta_T(t_{\mathrm{mat}}, \boldsymbol{q})| \frac{k_{\mathrm{B}} T_{(0)}(t)}{m_{\mathrm{H}} c^2} \leq \tfrac{3}{2}|\delta_T(t_{\mathrm{mat}}, \boldsymbol{q})| \frac{k_{\mathrm{B}} T_{(0)}(t_{\mathrm{dec}})}{m_{\mathrm{H}} c^2} = 4.10 \times 10^{-10} |\delta_T(t_{\mathrm{mat}}, \boldsymbol{q})|, \tag{322}$$

implying that

$$\delta_n(t, \boldsymbol{q}) \approx \delta_\varepsilon(t, \boldsymbol{q}), \quad t \geq t_{\mathrm{dec}}. \tag{323}$$

We take, however, $\delta_n(t_{\mathrm{mat}}, \boldsymbol{q}) \neq \delta_\varepsilon(t_{\mathrm{mat}}, \boldsymbol{q})$. The case $\delta_n(t_{\mathrm{mat}}, \boldsymbol{q}) = \delta_\varepsilon(t_{\mathrm{mat}}, \boldsymbol{q})$ has been discussed in Section XIII C 3 on the general solution of the evolution equation (298).

An initial value for the relative matter temperature perturbation $\delta_T(t_{\mathrm{mat}}, \boldsymbol{q})$ can be found as follows. A weak condition for growth is that a perturbation must lose some of its internal heat energy, i.e. $T_{(0)} s_{(1)}^{\mathrm{gi}} < 0$ or, equivalently, (285). A larger growth can be achieved if we take the initial values such that the heat loss of a density perturbation is larger than given by (309), i.e.

$$-k_{\mathrm{B}} T_{(0)}(t) \delta_n(t, \boldsymbol{q}) > \tfrac{1}{2} k_{\mathrm{B}} T_{(0)}(t) \bigl[3\delta_T(t_{\mathrm{mat}}, \boldsymbol{q}) - 2\delta_n(t, \boldsymbol{q})\bigr] \quad \Rightarrow \quad \delta_T(t_{\mathrm{mat}}, \boldsymbol{q}) < 0, \tag{324}$$

where we have also used the combined First and Second Laws of thermodynamics in the form (291). The criterion (324), yields a positive source term of the evolution equation (298) and, therefore, a larger growth than the homogeneous equation.

Finally, there are no observations of the growth rates of density perturbations. Therefore, we take the initial growth rates equal to zero, i.e.

$$\delta_\varepsilon'(t_{\mathrm{mat}}, \boldsymbol{q}) = \delta_n'(t_{\mathrm{mat}}, \boldsymbol{q}) = 0, \quad t_{\mathrm{dec}} \leq t_{\mathrm{mat}} \leq t_{\mathrm{p}}. \tag{325}$$

In the next two subsections we investigate the influence of both the particle mass and the initial time of star formation on the mass of a particular star. In Section XV we present the results of our new perturbation theory for structure formation.

B. The Influence of the Particle Mass on the Mass of a Star

Up till now we have assumed that the matter content of the universe after decoupling consists of protons. However, WMAP observations suggest that there may exist a considerable amount of, as yet unknown, dark matter (DM) particles. In this section we investigate the influence of the mean particle mass on the evolution of density perturbations. We assume that after decoupling, a gas consisting of particles with mean mass $\tilde{m}$, given by ($\alpha(t) > 0$)

$$\tilde{m} = \alpha(t_{\mathrm{mat}}) m_{\mathrm{H}}, \tag{326}$$

can be described by an equation of state of the form (264) with the mass m_{H} replaced by (326). As a crude estimate for $\tilde{m}$, we take the mean mass of the baryons and the DM particles, i.e.

$$\alpha(t) = \frac{\Omega_{\mathrm{bar}}(t) m_{\mathrm{H}} + \Omega_{\mathrm{dm}}(t) m_{\mathrm{dm}}}{\bigl[\Omega_{\mathrm{bar}}(t) + \Omega_{\mathrm{dm}}(t)\bigr] m_{\mathrm{H}}}, \tag{327}$$

where Ω_{bar} and Ω_{dm} are the baryon and dark matter particle densities respectively, given in units of the critical density, see the Friedmann equation (155).

By $\tilde{\mu}_{\mathrm{m}}$, we denote the parameter μ_{m} (315) in which m_{H} is replaced by $\alpha(t_{\mathrm{mat}}) m_{\mathrm{H}}$. Density perturbations in a gas with mean particle mass $\tilde{m}$ evolve in exactly the same way as perturbations in a gas of which the particles have mass m_{H} if $\mu_{\mathrm{m}} = \tilde{\mu}_{\mathrm{m}}$: in this case we have $\delta_\varepsilon(t, \boldsymbol{q}) = \tilde{\delta}_\varepsilon(t, \boldsymbol{q})$. From (315) it follows that

$$\tilde{\lambda}_{\mathrm{mat}} = \frac{\lambda_{\mathrm{mat}}}{\sqrt{\alpha(t_{\mathrm{mat}})}}, \tag{328}$$

where $\tilde{\lambda}_{\rm mat}$ is the scale of a density perturbation in a gas with mean particle mass $\tilde{m}$. Using (317), (326) and (328) we find

$$\tilde{M}(t_{\rm mat}) = \frac{M(t_{\rm mat})}{\sqrt{\alpha(t_{\rm mat})}}, \tag{329}$$

where $\tilde{M}(t_{\rm mat})$ refers to a perturbation mass in a gas with mean particle mass $\tilde{m}$. In the derivation of (329) we have assumed that $\tilde{n}_{(0)}(t_{\rm mat}) = n_{(0)}(t_{\rm mat})$, i.e. the particle number density is for a mean particle mass $\tilde{m}$ the same as for a particle mass $m_{\rm H}$. We thus have found that perturbations in a gas consisting of particles which are heavier (i.e. $\alpha(t_{\rm mat}) > 1$) than protons the collapse takes place at a smaller total mass than perturbations in ordinary matter. In other words, heavier particles yield lighter stars. We will illustrate (329) with two extreme cases, namely a mixture of heavy WIMPs (CDM) with a mass of $m_{\rm cdm} \approx 70\, m_{\rm H}$ [38] and baryons, and a mixture of ordinary matter and *hot dark matter* (HDM) with a mass of $m_{\rm hdm} = 1.5\,{\rm eV}/c^2$, as suggested by Nieuwenhuizen *et al.* [5, 40]. Using that $m_{\rm H} = 0.938\,{\rm GeV}/c^2$, we can calculate with the help of (157) and (327) the mean particle mass and, hence, α. Assuming that $\Omega_{\rm bar}(t)/\Omega_{\rm dm}(t) \approx \Omega_{\rm bar}(t_{\rm p})/\Omega_{\rm dm}(t_{\rm p})$ for both types of particles and for $t > t_{\rm dec}$, we find $\alpha(t) \approx \alpha(t_{\rm p})$, so that $\alpha_{\rm cdm}^{-1/2} \approx 0.13$ and $\alpha_{\rm hdm}^{-1/2} \approx 2.4$. This implies that if the dark matter particles are heavy WIMPs then the stars formed are much lighter than the stars formed in a universe filled with baryons only. On the other hand, if light neutrinos are the dark matter (HDM), then the stars formed will be heavier.

C. The Influence of Initial Time on the Mass of a Star

In this section we show that density fluctuations which start to contract at late times yield stars with smaller masses than early density perturbations. To show this, we consider the evolution equation (298).

It follows from equation (298) that perturbations starting to grow at $t_{\rm mat} > t_{\rm dec}$ and obeying the same initial conditions as fluctuations starting to contract at $t_{\rm dec}$, i.e.

$$\delta_\varepsilon(t_{\rm mat}, \boldsymbol{x}) = \delta_\varepsilon(t_{\rm dec}, \boldsymbol{x}), \quad \delta_\varepsilon'(t_{\rm mat}, \boldsymbol{x}) = \delta_\varepsilon'(t_{\rm dec}, \boldsymbol{x}), \quad \delta_T(t_{\rm mat}, \boldsymbol{x}) = \delta_T(t_{\rm dec}, \boldsymbol{x}), \tag{330}$$

evolve in exactly the same way, provided that

$$\mu_{\rm m}(t_{\rm mat}) = \mu_{\rm m}(t_{\rm dec}), \tag{331}$$

where we have written $\mu_{\rm m}(t_0)$ to denote the value of the parameter $\mu_{\rm m}$ taken at time $t = t_0$. Using the relation (331) we can relate the masses of density perturbations starting at different times. Replacing the initial time $t_{\rm mat}$ in equation (298) by the initial time $t_{\rm dec}$, we get for the parameter $\mu_{\rm m}$, (315), at $t = t_{\rm dec}$

$$\mu_{\rm m}(t_{\rm dec}) = \frac{2\pi}{\lambda_{\rm dec}} \frac{1}{\mathcal{H}(t_{\rm p})\big[z(t_{\rm dec}) + 1\big]^{3/2}} \sqrt{\frac{5}{3}\frac{k_{\rm B}T_{(0)}(t_{\rm dec})}{m_{\rm H}}}. \tag{332}$$

Equating (315) and (332), we arrive at

$$\lambda_{\rm mat} = \lambda_{\rm dec}\sqrt{\frac{z(t_{\rm dec}) + 1}{z(t_{\rm mat}) + 1}}, \tag{333}$$

i.e. the evolution of a density perturbation starting at $t = t_{\rm dec}$ with scale $\lambda_{\rm dec}$ is exactly equal to the evolution of a density perturbation starting at $t = t_{\rm mat}$ with scale $\lambda_{\rm mat}$, if and only if the relation between the scales obeys (333). Since $n_{(0)} \propto a^{-3} \propto (z+1)^3$, we have

$$n_{(0)}(t_{\rm mat}) = n_{(0)}(t_{\rm dec}) \left(\frac{z(t_{\rm mat}) + 1}{z(t_{\rm dec}) + 1}\right)^3. \tag{334}$$

Using these expressions and (317), we finally arrive at

$$M(t_{\rm mat}) = M(t_{\rm dec}) \left(\frac{z(t_{\rm mat}) + 1}{z(t_{\rm dec}) + 1}\right)^{\frac{3}{2}}. \tag{335}$$

This expression relates the masses of density perturbations, which start to evolve at different times $t_{\rm dec}$ and $t_{\rm mat}$ with the same initial values (330) and (331). Consequently, stars formed from late time fluctuations have smaller masses.

XV. STAR FORMATION: RESULTS

The standard cosmological theory of small perturbations is characterized by $\delta_n = \delta_\varepsilon$, (308). This is, however, too restrictive. Although $|\delta_n - \delta_\varepsilon|$ is very small, it need not be zero, as follows from (322). Since $m_{\rm H}c^2 \gg k_{\rm B}T_{(0)}$, it follows from (282) and (286) that the quantities δ_T, δ_p and δ_n are not, beforehand, confined to small values: small changes in $\Delta \equiv \delta_n - \delta_\varepsilon$ may lead to large changes in δ_T, δ_p and δ_n. These quantities must only fulfill the linearity conditions $|\delta_T| \leq 1$, $|\delta_p| \leq 1$ and $|\delta_n| \leq 1$. The fact that δ_n need not be exactly equal to δ_ε is paramount for the formation of structure in the universe.

In this section we will solve the evolution equation (298) numerically. To that end we use the differential equation solver `lsodar` with root finding capabilities, included in the package `deSolve`, which, in turn, is included in R, a system for statistical computation and graphics [41]. We investigate star formation which starts at cosmological redshifts $z = 1091$ and $z = 1$.

A. Star Formation starting at $z = 1091$: Population III Stars

At the moment of decoupling of matter and radiation photons could not ionize matter any more and the two constituents fell out of thermal equilibrium. As a consequence, the pressure drops from a very high radiation pressure $p = \frac{1}{3}a_{\rm B}T_\gamma^4$ just before decoupling to a very low gas pressure $p = nk_{\rm B}T$ after decoupling. This fast transition from a high pressure epoch to a very low pressure era may result in large relative pressure perturbations. In this subsection we study the influence of these relative pressure perturbations on the formation of stars.

1. *Initial Values*

In order to integrate equation (298), we need initial values. The initial value $\delta_\varepsilon(t_{\rm dec}, \boldsymbol{q})$ is related to the relative perturbation $\delta_{T_\gamma}(t_{\rm dec}, \boldsymbol{q})$ in the *background radiation*. Using (205) and (261), we get at the end of the plasma era

$$\delta_{T_\gamma}(t_{\rm dec}, \boldsymbol{q}) = \frac{1}{4}\left[\frac{n_{(0)}(t_{\rm dec})m_{\rm H}c^2}{a_{\rm B}T^4_{(0)\gamma}(t_{\rm dec})}\bigl[\delta_\varepsilon(t_{\rm dec}, \boldsymbol{q}) - \delta_n(t_{\rm dec}, \boldsymbol{q})\bigr] + \delta_\varepsilon(t_{\rm dec}, \boldsymbol{q})\right]. \tag{336}$$

Using that $n_{(0)} \propto a^{-3}$ and $T_{(0)\gamma} \propto a^{-1}$, we get from (232) and (318)

$$n_{(0)}(t_{\rm dec})m_{\rm H}c^2 = a_{\rm B}T^4_{(0)\gamma}(t_{\rm dec})\frac{z(t_{\rm eq}) + 1}{z(t_{\rm dec}) + 1}, \tag{337}$$

so that (336) can be rewritten as

$$\delta_{T_\gamma}(t_{\rm dec}, \boldsymbol{q}) = \frac{1}{4}\left[\frac{z(t_{\rm eq}) + 1}{z(t_{\rm dec}) + 1}\bigl[\delta_\varepsilon(t_{\rm dec}, \boldsymbol{q}) - \delta_n(t_{\rm dec}, \boldsymbol{q})\bigr] + \delta_\varepsilon(t_{\rm dec}, \boldsymbol{q})\right]. \tag{338}$$

In order to eliminate $\delta_n(t_{\rm dec}, \boldsymbol{q})$, we use that at the end of the plasma era we have, according to (263),

$$\delta_n(t_{\rm dec}, \boldsymbol{q}) - \frac{\delta_\varepsilon(t_{\rm dec}, \boldsymbol{q})}{1 + w(t_{\rm dec})} \approx 0, \tag{339}$$

where, using (261) and (337),

$$w(t_{\rm dec}) \equiv \frac{p_{(0)}(t_{\rm dec})}{\varepsilon_{(0)}(t_{\rm dec})} = \frac{z(t_{\rm dec}) + 1}{3z(t_{\rm eq}) + 3z(t_{\rm dec}) + 6} \approx 0.085. \tag{340}$$

Combining (339) and (340), we find

$$\delta_n(t_{\rm dec}, \boldsymbol{q}) = \delta_\varepsilon(t_{\rm dec}, \boldsymbol{q})\frac{3z(t_{\rm eq}) + 3z(t_{\rm dec}) + 6}{3z(t_{\rm eq}) + 4z(t_{\rm dec}) + 7} \approx 0.92\ \delta_\varepsilon(t_{\rm dec}, \boldsymbol{q}). \tag{341}$$

We can now rewrite (338) as

$$\delta_{T_\gamma}(t_{\rm dec}, \boldsymbol{q}) = \delta_\varepsilon(t_{\rm dec}, \boldsymbol{q})\frac{z(t_{\rm eq}) + z(t_{\rm dec}) + 2}{3z(t_{\rm eq}) + 4z(t_{\rm dec}) + 7} \approx 0.31\ \delta_\varepsilon(t_{\rm dec}, \boldsymbol{q}). \tag{342}$$

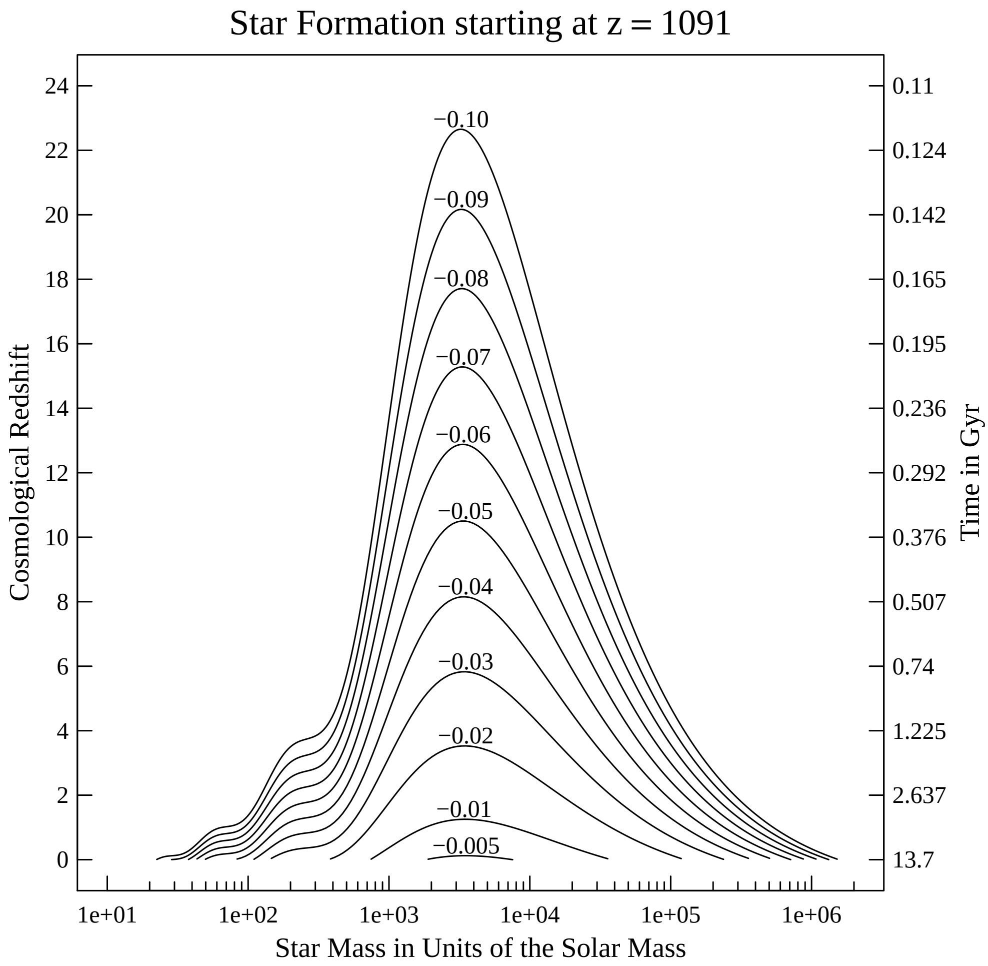

Figure 1: The curves give the redshift at which a linear perturbation in the particle number density with initial values $\delta_n(t_{\rm dec}, \boldsymbol{q}) \approx 10^{-5}$ and $\delta_n'(t_{\rm dec}, \boldsymbol{q}) = 0$ starting to grow at an initial redshift of $z(t_{\rm dec}) = 1091$ becomes non-linear, i.e. $\delta_n \approx 1$. During the evolution we have $\delta_p(t, \boldsymbol{q}) = \delta_T(t_{\rm dec}, \boldsymbol{q}) + \delta_n(t, \boldsymbol{q})$. The numbers at each of the curves are the initial relative perturbations in the matter temperature $\delta_T(t_{\rm dec}, \boldsymbol{q})$. For each curve, the maximum is at $3.4 \times 10^3\,\mathrm{M}_\odot$. If CDM is present then the peak mass is at $4.5 \times 10^2\,\mathrm{M}_\odot$.

From the WMAP observation (231f) we find, using (342),

$$\delta_\varepsilon(t_{\rm dec}, \boldsymbol{q}) \approx 3.3 \times 10^{-5}, \quad \delta_\varepsilon'(t_{\rm dec}, \boldsymbol{q}) \approx 0, \tag{343}$$

where we have assumed, for lack of observations, that during the decoupling of matter and radiation the growth rate δ_ε' is very small, see (325). With (343) the condition (323) implies that

$$\delta_n(t_{\rm dec}, \boldsymbol{q}) \approx 3.3 \times 10^{-5}, \quad \delta_n'(t_{\rm dec}, \boldsymbol{q}) \approx 0, \tag{344}$$

i.e. WMAP observations demand that, just after decoupling, also the relative particle number density perturbations are very small.

As we have shown, see (288), the relative matter temperature perturbation δ_T is very nearly constant for a contracting or expanding density perturbation. From the fact that the pressure is given by $p = nk_{\rm B}T$ and (344) it follows that large pressure perturbations can, just after decoupling, only be realized by large matter temperature perturbations. In our calculations we take, therefore, $|\delta_T(t_{\rm dec}, \boldsymbol{x})|$ in the range 0.5%–10%. With (287) and (344) this implies that, initially,

$$\delta_p(t_{\rm dec}, \boldsymbol{q}) \approx \delta_T(t_{\rm dec}, \boldsymbol{q}). \tag{345}$$

In other words, the perturbation in the pressure is, initially, mainly determined by a perturbation in the matter temperature. In view of (324) and (344) we have to take for the initial value of the relative matter temperature perturbation

$$\delta_T(t_{\rm dec}, \boldsymbol{q}) < 0, \tag{346}$$

in order to find growing density perturbations. We now have gathered the necessary ingredients to integrate the evolution equation (298) numerically.

2. Results

Figure 1 has been constructed [41] as follows. For each choice of $\delta_T(t_{\rm dec}, \boldsymbol{q})$ we integrate equation (298) for a large number of values for $\lambda_{\rm dec}$, using the initial values (343) and (344). The integration starts at $\tau \equiv t/t_{\rm dec} = 1$, i.e. at $z = z(t_{\rm dec})$ and will be halted if either $z = 0$ [i.e. $\tau = \tau_{\rm end}$, (300)] or $\delta_\varepsilon(t, \boldsymbol{q}) = 1$ has been reached. One integration run yields one point on the curve for a particular choice of $\lambda_{\rm dec}$ if $\delta_\varepsilon(t, \boldsymbol{q}) = 1$ has been reached for $z > 0$. If the integration halts at $z = 0$ and $\delta_\varepsilon(t, \boldsymbol{q}) < 1$, then the perturbation belonging to that particular $\lambda_{\rm dec}$ has not yet reached its non-linear phase today, i.e. at $t = 13.7\,{\rm Gyr}$. On the other hand, if the integration is stopped at $\delta_\varepsilon(t, \boldsymbol{q}) = 1$, then the perturbation has become non-linear within 13.7 Gyr. In Figure 1 we have used the star mass $M(t_{\rm dec})$, expressed in solar masses, instead of the scale $\lambda_{\rm dec}$ of a density perturbation. To that end we have used expression (321).

The above described procedure is repeated for $\delta_T(t_{\rm dec}, \boldsymbol{q})$ in the range $-0.005, -0.01, -0.02, \dots, -0.1$. During the evolution, the relative pressure perturbation evolves as

$$\delta_p(t, \boldsymbol{q}) = \delta_T(t_{\rm dec}, \boldsymbol{q}) + \delta_n(t, \boldsymbol{q}). \tag{347}$$

The fastest growth is seen for perturbations with a mass of approximately $3.4 \times 10^3\,{\rm M}_\odot$. This value is nearly independent of the initial value of the matter temperature perturbation $\delta_T(t_{\rm dec}, \boldsymbol{q})$. Even density perturbations with an initial matter temperature perturbation as small as $\delta_T(t_{\rm dec}, \boldsymbol{q}) = -0.5\%$ reach their non-linear phase at $z = 0.13$ ($T_{(0)\gamma} = 3.1\,{\rm K}$, $t = 11.5\,{\rm Gyr}$) provided that its mass is around $3.4 \times 10^3\,{\rm M}_\odot$. Perturbations with masses smaller than $3.4 \times 10^3\,{\rm M}_\odot$ reach their non-linear phase at a later time, because their internal gravity is weaker. On the other hand, perturbations with masses larger than $3.4 \times 10^3\,{\rm M}_\odot$ have larger scales so that they cool down slower, resulting also in a smaller growth rate. Since the growth rate decreases rapidly for perturbations with masses below $3.4 \times 10^3\,{\rm M}_\odot$, the latter may be considered as the relativistic counterpart of the *Jeans mass*.

Figure 1 has been calculated for a baryonic cosmic fluid, without CDM or HDM. If CDM is present, then it follows from (329) that the graphs shift towards smaller masses such that the Jeans masses are at $4.5 \times 10^2\,{\rm M}_\odot$. For HDM the peaks are at $8.2 \times 10^3\,{\rm M}_\odot$.

Finally, using (321), we find that all star masses in Figure 1 start to contract at decoupling from density perturbations with diameters less than 52 pc, which is much smaller than the particle horizon size (Appendix D) $d_{\rm H}(t_{\rm dec}) = 349\,{\rm kpc}$.

B. Star Formation starting at $z = 1$

At $z = 1$ the interstellar gas has been diluted so much that star formation is not possible anymore in the interstellar gas with only the negative matter temperature perturbation $\delta_T(t_{\rm mat}, \boldsymbol{q})$ as the driving force. Consequently, late time star formation can only take place in regions which have a higher density with respect to the intergalactic space. Therefore, late time star formation takes place mainly within galaxies. As can be seen from Figures 2 and 3, the initial density perturbation should be at least of the order of $\delta_n \approx \delta_\varepsilon \approx 0.70$, in order to yield eventually a gravitational collapse for $z > 0$.

We have considered two extreme cases of star formation starting at $z(t_{\rm mat}) = 1$ or, equivalently, $t_{\rm mat} \approx 4.8\,{\rm Gyr}$. In both cases, we assume that the initial growth rate vanishes, (325). In the first case, depicted in Figure 2, we have chosen $\delta_T(t_{\rm mat}, \boldsymbol{q}) = -1$, the smallest value that a relative perturbation can have without violating the linearity conditions. During the evolution of a density perturbation, we have, according to (289),

$$\delta_p(t, \boldsymbol{q}) = -1 + \delta_n(t, \boldsymbol{q}), \tag{348}$$

so that during the evolution the relative pressure perturbation is negative. This is the most favorable situation for a perturbation to grow: pressure cooperates with gravitation. The most conspicuous feature in Figure 2 is the sharp lower limit of star formation. Density perturbations with masses below $0.2\,{\rm M}_\odot$, do not become non-linear before

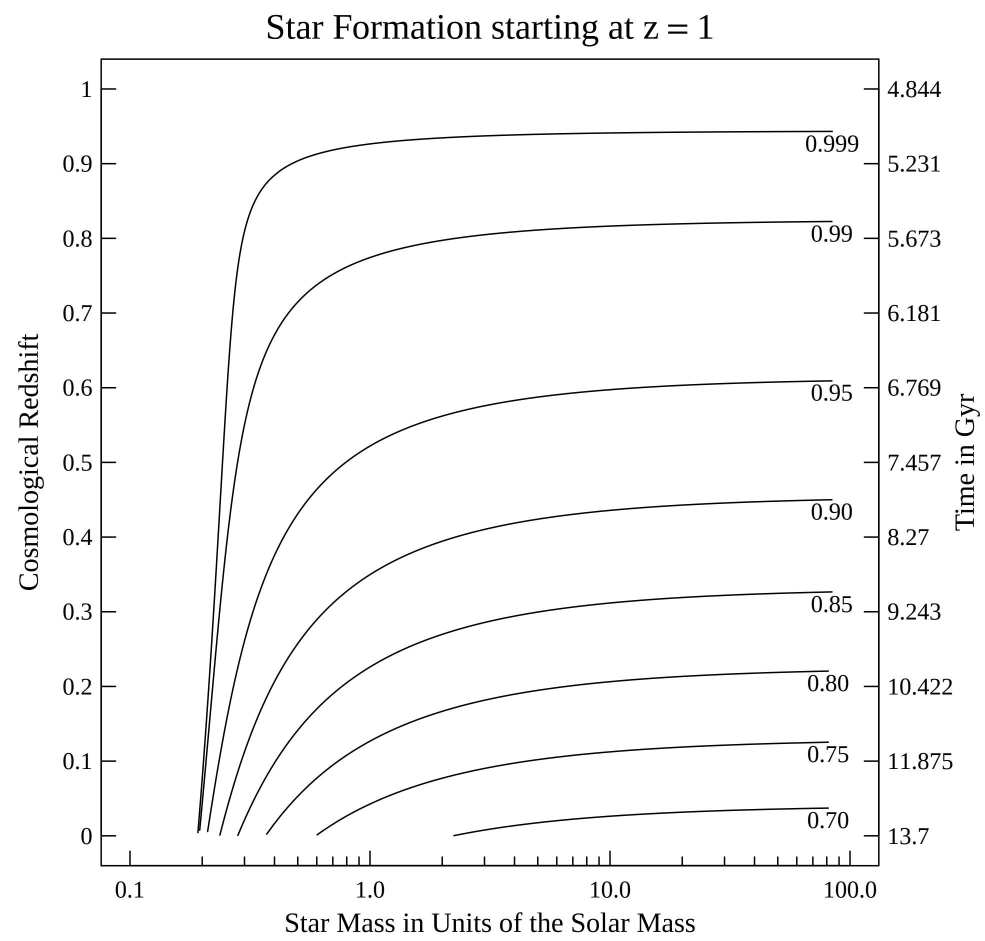

Figure 2: Star formation for $\delta_T(t, \boldsymbol{q}) = -1$. The curves give the redshift at which a linear density perturbation starting to grow at an initial redshift of $z(t_{\mathrm{mat}}) = 1$ becomes non-linear, i.e. $\delta_n \approx \delta_\varepsilon \approx 1$. The numbers at each of the curves are the initial relative density perturbations $\delta_n(t_{\mathrm{mat}}, \boldsymbol{q}) \approx \delta_\varepsilon(t_{\mathrm{mat}}, \boldsymbol{q})$. During the evolution we have $\delta_p(t, \boldsymbol{q}) = -1 + \delta_n(t, \boldsymbol{q})$. For masses in excess of $100\,\mathrm{M}_\odot$, the growth becomes independent of the scale of a perturbation, i.e. the growth is proportional to $t^{2/3}$.

$z = 0$ is reached. Apparently, the gravitational field is, for masses below $0.2\,\mathrm{M}_\odot$, too weak to collapse in due time and, eventually, become a star. Another notable characteristic of star formation is that it is slow: even if the initial density perturbation is as high as $\delta_n(t_{\mathrm{mat}}, \boldsymbol{q}) = 0.999$ it takes, for a $100\,\mathrm{M}_\odot$ perturbation, approximately $214\,\mathrm{Myr}$ to reach the value $\delta_n(t, \boldsymbol{q}) = 1$.

The second case, characterized by $\delta_T(t_{\mathrm{mat}}, \boldsymbol{q}) = 0$ and summarized in Figure 3, is the most detrimental for star formation, because during the evolution of a density perturbation we have

$$\delta_p(t, \boldsymbol{q}) = \delta_n(t, \boldsymbol{q}), \tag{349}$$

so that the pressure opposes the contraction of a density perturbation. However, if the internal gravity of a perturbation is strong enough, then gravity will overcome the pressure and the perturbation will eventually collapse to form a star. Due to the counteracting pressure, the lower limit of star formation is much larger than in the case of cooperating pressure, namely $0.8\,\mathrm{M}_\odot$.

Both extreme cases have one common characteristic. For perturbations with masses larger than $100\,\mathrm{M}_\odot$, the collapse time is nearly the same, as can be seen from Figures 2 and 3. Apparently, the internal gravitational field is so strong that opposing or cooperating pressure perturbations do not play a role anymore.

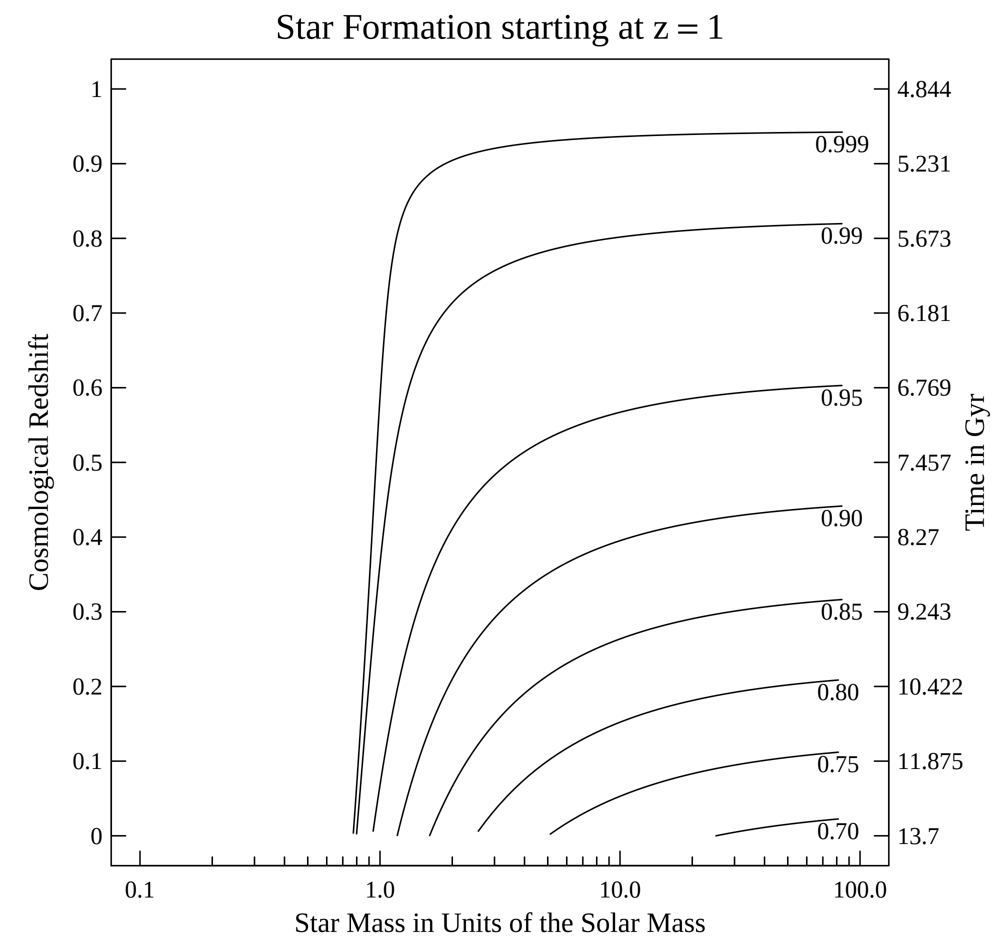

Figure 3: Star formation for $\delta_T(t, \boldsymbol{q}) = 0$. The curves give the redshift at which a linear density perturbation starting to grow at an initial redshift of $z(t_{\rm mat}) = 1$ becomes non-linear, i.e. $\delta_n = \delta_\varepsilon \approx 1$. The numbers at each of the curves are the initial relative perturbations $\delta_n(t_{\rm mat}, \boldsymbol{q}) = \delta_\varepsilon(t_{\rm mat}, \boldsymbol{q})$. During the evolution we have $\delta_p(t, \boldsymbol{q}) = \delta_n(t, \boldsymbol{q}) = \delta_\varepsilon(t, \boldsymbol{q})$. For masses in excess of $100\,\mathrm{M}_\odot$, the growth becomes independent of the scale of a perturbation, i.e. the growth is proportional to $t^{2/3}$.

XVI. STANDARD NEWTONIAN THEORY OF COSMOLOGICAL PERTURBATIONS

The new evolution equations (245a) and (293) are different from their standard counterparts (350) and (357) respectively. In this section we explain the differences and we show why equations (350) and (357) should not be used anymore in the study of cosmological density perturbations.

Padmanabhan had already in 1993 the supposition that the Newtonian theory of cosmological density perturbations is questionable. On page 136 of his textbook [39] he states:

> To avoid any misunderstanding, we emphasize the following fact: it is not possible to study cosmology using Newtonian gravity without invoking procedures which are mathematically dubious.

Padmanabhan does not conclude, however, that the Newtonian theory of cosmological perturbations is incorrect, since he subsequently states:

> However, if we are only interested in regions much smaller than the characteristic length scale set by the curvature of space-time, then one can introduce a valid approximation to general relativity.

Up till now, this point of view is considered as standard knowledge. This is due to the fact that the standard equation (357) of the Newtonian theory for small-scale perturbations is, in the low velocity limit $v_{\rm s}/c \to 0$, similar to the

relativistic equation (362) derived from the General Theory of Relativity. This similarity has led to much confusion in the literature (see, for example, [32–34, 42, 43]), as we will now explain. First, we remark that gauge problems, although time and space coordinates may be chosen freely according to (220), do *not* occur in the Newtonian theory of gravity. This is because in the Newtonian theory itself the universe is *static*, i.e. $\dot{\varepsilon}_{(0)} = 0$ and $\dot{n}_{(0)} = 0$, implying with (219) that both $\varepsilon_{(1)}$ and $n_{(1)}$ are independent of the choice of a system of reference. Now, the reasoning in the literature is as follows. The quantity $\delta \equiv \varepsilon_{(1)}/\varepsilon_{(0)}$ occurring in the Newtonian equation (357) is, according to the standard knowledge, independent of the gauge choice, since this equation is derived from the Newtonian theory, in which gauge problems concerning perturbed scalar quantities do not occur. Furthermore, the Newtonian equation (357) is valid only for density perturbations with scales smaller than the horizon size. By virtue of the resemblance of the Newtonian equation (357) and the relativistic equation (362) it is, therefore, put forward that

> a gauge dependent quantity, such as $\delta \equiv \varepsilon_{(1)}/\varepsilon_{(0)}$, which is initially larger than the horizon, becomes automatically gauge-invariant as soon as the perturbation becomes smaller than the horizon.

This viewpoint is, however, incorrect as we will now show in detail. Firstly, we remark that the universe *cannot be static* in the non-relativistic *limit* (see Section XII) of the General Theory of Relativity, since $H \to 0$ violates the Einstein equations (214). Secondly, it has been shown in Section XII that there remains some gauge freedom in the non-relativistic limit, namely the freedom to shift time coordinates, $x^0 \to x^0 - \psi$, and the freedom to choose spatial coordinates, $x^i \to x^i - \chi^i(\boldsymbol{x})$. This coordinate freedom is a well-known and natural property of Newtonian physics and, therefore, should follow from a relativistic perturbation theory in the low velocity limit $v_{\mathrm{s}}/c \to 0$. Thirdly, we will show that, just because of the resemblance of the relativistic and Newtonian equation, the gauge function ψ also occurs in the solution (359) of the Newtonian equation (357). Consequently, the standard equation of the Newtonian theory has no physical significance: due to the appearance of the gauge function ψ in one of the two independent solutions, one cannot impose initial conditions to arrive at a physical solution. From the fact that, in the non-relativistic limit of the General Theory of Relativity, the universe is not static, combined with the freedom in time coordinates implies that if a quantity, such as $\delta \equiv \varepsilon_{(1)}/\varepsilon_{(0)}$, is gauge dependent in the General Theory of Relativity it is also gauge dependent in the non-relativistic limit, as follows from (219). The important conclusion is, therefore, that *a quantity which has no physical significance outside the horizon, does not become a physical quantity inside the horizon.* This conclusion is consistent with the facts that, in the non-relativistic limit of our perturbation theory, the gauge-invariant quantities $\varepsilon^{\mathrm{gi}}_{(1)}$, (225), and $n^{\mathrm{gi}}_{(1)}$, (228), survive, while the gauge dependent quantities $\varepsilon_{(1)}$ and $n_{(1)}$ disappear from the scene, as we have shown in detail in Section XII. Hence, with $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$ surviving in the non-relativistic limit, there is indeed no gauge problem in the Newtonian theory of gravity.

The fact that linear perturbation theory is plagued by the gauge solutions (356) and (363) has already been pointed out by Lifschitz [6, 22, 23] as early as 1946. In 1980, thirty-four years later, Press and Vishniac [11] called attention to the same issue. In spite of these warnings, the standard equations (350) and (357) are still ubiquitous in the cosmological literature. Apparently, the cosmological gauge problem is quite persevering.

In the next two subsections, we will elucidate the gauge problem. We go one step further than Lifschitz and Press and Vishniac. These researchers have shown that only for large scales the solutions of the standard equations are infected by spurious gauge modes. We show, in addition, that the solutions (352) and (359) of the standard equations (350) and (357) contain the gauge function $\psi(\boldsymbol{x})$, *independent* of the scale of a perturbation. Consequently, the important conclusion must be that

> the Newtonian theory of gravity is *not* suitable to study the evolution of cosmological density perturbations.

In order to show that the Newtonian theory of cosmological perturbations is invalid, we consider a flat ($k = 0$) FLRW universe with a vanishing cosmological constant ($\Lambda = 0$) in the radiation-dominated era and the era after decoupling of matter and radiation.

A. Radiation-dominated Universe

The standard equation for the density contrast function δ which can be found, for example, in the textbook of Peacock [44], equation (15.25), is given by

$$\ddot{\delta} + 2H\dot{\delta} - \left(\frac{1}{3}\frac{\nabla^2}{a^2} + \tfrac{4}{3}\kappa\varepsilon_{(0)}\right)\delta = 0. \tag{350}$$

Compare this equation with equation (245a) for the gauge-invariant contrast function δ_ε. Equation (350) is derived by using special relativistic fluid mechanics and the Newtonian theory of gravity with a relativistic source term. In agreement with the text under equation (15.25) of this textbook, the term $-\frac{1}{3}\nabla^2\delta/a^2$ has been added. The same

result, equation (350), can be found in Weinberg's classic [35], equation (15.10.57) with $p = \frac{1}{3}\rho$ and $v_{\mathrm{s}} = 1/\sqrt{3}$. Note, that equation (350) cannot be derived from the General Theory of Relativity.

Using (240), (241), (248), (251) and (252), we can rewrite equation (350) in the form

$$\delta'' + \frac{1}{\tau}\delta' + \left(\frac{\mu_{\mathrm{r}}^2}{4\tau} - \frac{1}{\tau^2}\right)\delta = 0, \tag{351}$$

where the constant μ_{r} is given by (254). A prime denotes differentiation with respect to τ, (251). The general solution of this equation is found to be

$$\delta(\tau, \boldsymbol{q}) = \frac{8C_1(\boldsymbol{q})}{\mu_{\mathrm{r}}^2} J_2\big(\mu_{\mathrm{r}}\sqrt{\tau}\big) + \psi(\boldsymbol{q})\pi\mu_{\mathrm{r}}^2 H(t_{\mathrm{rad}}) Y_2\big(\mu_{\mathrm{r}}\sqrt{\tau}\big), \tag{352}$$

where $C_1(\boldsymbol{q})$ and $\psi(\boldsymbol{q})$ are arbitrary functions (the integration 'constants') and $J_\nu(x)$ and $Y_\nu(x)$ are Bessel functions of the first and second kind respectively. The factors $8/\mu_{\mathrm{r}}^2$ and $\pi\mu_{\mathrm{r}}^2 H(t_{\mathrm{rad}})$ have been inserted for convenience. Thus, the standard equation (350) yields oscillating density perturbations with a *decaying* amplitude.

For large-scale perturbations ($|\boldsymbol{q}| \to 0$ or, equivalently, $\mu_{\mathrm{r}} \to 0$), the asymptotic expressions for the Bessel functions J_2 and Y_2 are given by

$$J_2\big(\mu_{\mathrm{r}}\sqrt{\tau}\big) \approx \frac{\mu_{\mathrm{r}}^2}{8}\tau, \quad Y_2\big(\mu_{\mathrm{r}}\sqrt{\tau}\big) \approx -\frac{4}{\pi\mu_{\mathrm{r}}^2}\tau^{-1}. \tag{353}$$

Substituting these expressions into (352), it is found for large-scale perturbations that

$$\delta(\tau) = C_1\tau - 4H(t_{\mathrm{rad}})\psi\tau^{-1}, \tag{354}$$

where we have used that $C_1(|\boldsymbol{q}| \to 0) = C_1$ and $\psi(|\boldsymbol{q}| \to 0) = \psi$ become constants in the large-scale limit. Large-scale perturbations can also be obtained from the standard equation (350) by substituting $\nabla^2\delta = 0$, i.e.

$$\ddot{\delta} + 2H\dot{\delta} - \tfrac{4}{3}\kappa\varepsilon_{(0)}\delta = 0. \tag{355}$$

The general solution of this equation is, using (351) with $\mu_{\mathrm{r}} = 0$, given by (354). Thus far, the functions $C_1(\boldsymbol{q})$ and $\psi(\boldsymbol{q})$ are the integration 'constants' which can be determined by the initial values $\delta(t_{\mathrm{rad}}, \boldsymbol{q})$ and $\dot{\delta}(t_{\mathrm{rad}}, \boldsymbol{q})$.

However, equation (355) can, in contrast to equation (350), also be derived from the General Theory of Relativity: see the derivation in Appendix E. As a consequence, equation (355) is found to be also a *relativistic* equation, implying that the quantity $\delta = \varepsilon_{(1)}/\varepsilon_{(0)}$ is gauge dependent. Therefore, the second term in the solution (354) is not a physical mode, but equal to the gauge mode

$$\delta_{\mathrm{gauge}}(\tau) = \psi\frac{\dot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} = -4H(t_{\mathrm{rad}})\psi\tau^{-1}, \tag{356}$$

as follows from (2a), (239b) and (240a). Consequently, the constant ψ in (354) and, hence, the function $\psi(\boldsymbol{q})$ in (352) should *not* be interpreted as an integration constant, but as a gauge function, which cannot be determined by imposing initial value conditions, see Appendix C for a detailed explanation. Thus, the general solution (352) of the standard equation (350) depends on the gauge function $\psi(\boldsymbol{q})$ and has, as a consequence, no physical significance. This, in turn, implies that the standard equation (350) does *not* describe the evolution of density perturbations.

Here the negative effect of the gauge function is clearly seen: as yet it was commonly accepted that small-scale perturbations in the radiation-dominated era of a flat FLRW universe oscillate with a *decaying* amplitude, according to (352). The treatise presented in this article reveals, however, that small-scale density perturbations oscillate with an *increasing* amplitude, according to (260). This is the real behavior of a small-scale density perturbation.

B. Era after Decoupling of Matter and Radiation

The standard perturbation equation of the Newtonian theory of gravity is derived from *approximate*, *non-relativistic* equations. It reads

$$\ddot{\delta} + 2H\dot{\delta} - \left(\frac{v_{\mathrm{s}}^2}{c^2}\frac{\nabla^2}{a^2} + \tfrac{1}{2}\kappa\varepsilon_{(0)}\right)\delta = 0, \tag{357}$$

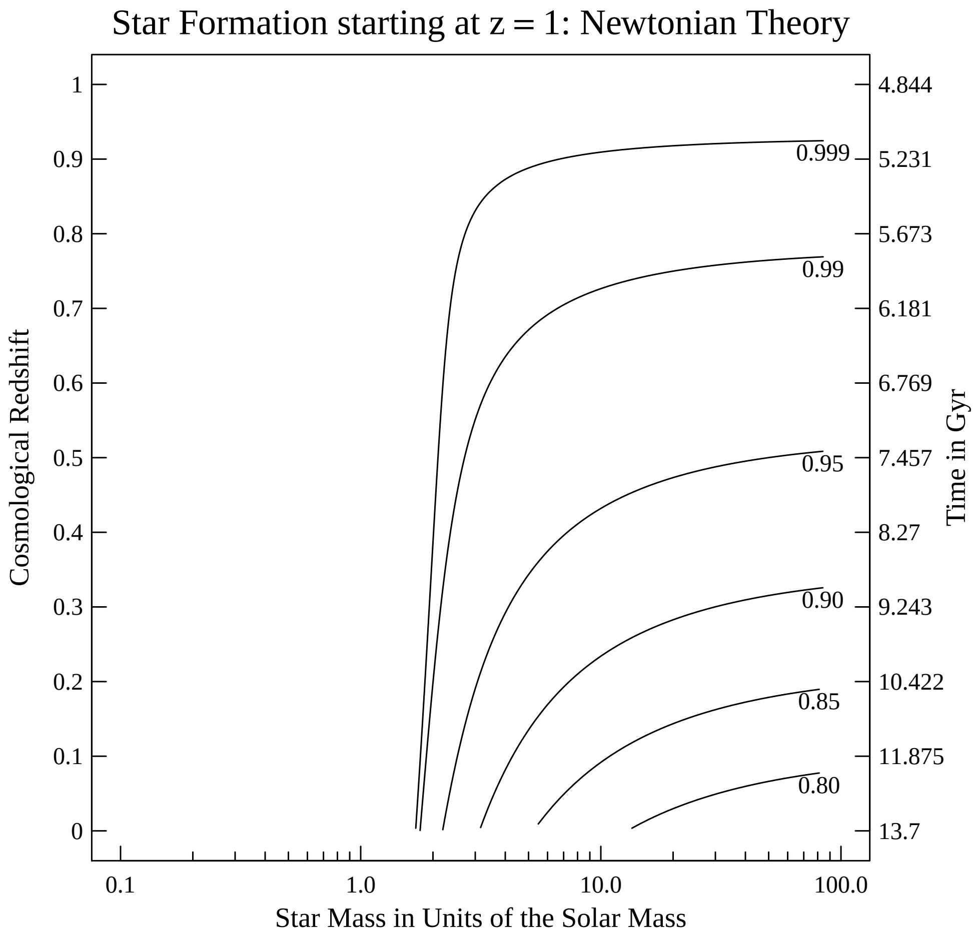

Figure 4: Star formation according to the standard equation (357) with $\delta_T(t,\boldsymbol{q})=0$. The curves give the redshift at which a linear density perturbation starting to grow at an initial redshift of $z(t_{\text{mat}})=1$ becomes non-linear, i.e. $\delta_n=\delta_\varepsilon=\delta_p\approx 1$. The numbers at each of the curves are the initial relative density perturbations $\delta_n(t_{\text{mat}},\boldsymbol{q})=\delta_\varepsilon(t_{\text{mat}},\boldsymbol{q})=\delta_p(t_{\text{mat}},\boldsymbol{q})$. During the evolution we have $\delta_p(t,\boldsymbol{q})=\delta_n(t,\boldsymbol{q})=\delta_\varepsilon(t,\boldsymbol{q})$. For masses in excess of $100\,\mathrm{M}_\odot$, the growth becomes independent of the scale of a perturbation, i.e. the growth is proportional to $t^{2/3}$.

where v_{s} is the speed of sound. (See, for example, Weinberg [35], Section 15.9, or Peacock [44], Section 15.2.) Compare this equation with equation (293) for the gauge-invariant contrast function δ_ε. Just as equation (350), the standard equation (357) cannot be derived from Einstein's gravitation theory.

Using (248), (268), (269), (296) and (297), equation (357) can rewritten in the form

$$\delta''+\frac{4}{3\tau}\delta'+\left(\frac{4}{9}\frac{\mu_{\text{m}}^2}{\tau^{8/3}}-\frac{2}{3\tau^2}\right)\delta=0, \tag{358}$$

where the constant μ_{m} is given by (295). A prime denotes differentiation with respect to τ, (296). The general solution of equation (358) is found to be

$$\delta(\tau,\boldsymbol{q})=\left[\frac{4}{3}D_1(\boldsymbol{q})\sqrt{\pi\mu_{\text{m}}^5}J_{-\frac{5}{2}}\big(2\mu_{\text{m}}\tau^{-1/3}\big)-\frac{45}{8}\psi(\boldsymbol{q})H(t_{\text{mat}})\sqrt{\frac{\pi}{\mu_{\text{m}}^5}}J_{+\frac{5}{2}}\big(2\mu_{\text{m}}\tau^{-1/3}\big)\right]\tau^{-1/6}, \tag{359}$$

where $D_1(\boldsymbol{q})$ and $\psi(\boldsymbol{q})$ are arbitrary functions (the 'constants' of integration) and $J_{\pm\nu}(x)$ is the Bessel function of the first kind. The factors $\frac{4}{3}\sqrt{\pi\mu_{\text{m}}^5}$ and $\frac{45}{8}H(t_{\text{mat}})\sqrt{\pi/\mu_{\text{m}}^5}$ have been inserted for convenience.

We now consider large-scale perturbations characterized by $\nabla^2\delta = 0$ (i.e. $|\boldsymbol{q}| \to 0$) or perturbations of all scales in the limit $v_{\mathrm{s}}/c \to 0$. Both limits imply $\mu_{\mathrm{m}} \to 0$, as follows from (295). The asymptotic expressions for the Bessel functions in the limit $\mu_{\mathrm{m}} \to 0$ are given by

$$J_{-\frac{5}{2}}\left(2\mu_{\mathrm{m}}\tau^{-1/3}\right) \approx \frac{3}{4\sqrt{\pi\mu_{\mathrm{m}}^5}}\tau^{5/6}, \quad J_{+\frac{5}{2}}\left(2\mu_{\mathrm{m}}\tau^{-1/3}\right) \approx \frac{8}{15}\sqrt{\frac{\mu_{\mathrm{m}}^5}{\pi}}\tau^{-5/6}. \tag{360}$$

Substituting these expressions into the general solution (359), results in

$$\delta(\tau) = D_1\tau^{2/3} - 3H(t_{\mathrm{mat}})\psi\tau^{-1}. \tag{361}$$

where we have used that $D_1(|\boldsymbol{q}| \to 0) = D_1$ and $\psi(|\boldsymbol{q}| \to 0) = \psi$ become constants in either the large-scale limit $|\boldsymbol{q}| \to 0$ or in the limit $v_{\mathrm{s}}/c \to 0$. In the limit $\mu_{\mathrm{m}} \to 0$, equation (357) reduces to

$$\ddot{\delta} + 2H\dot{\delta} - \tfrac{1}{2}\kappa\varepsilon_{(0)}\delta = 0. \tag{362}$$

Using (358) with $\mu_{\mathrm{m}} = 0$, we find that the general solution of equation (362) is given by (361). Thus far, the functions $D_1(\boldsymbol{q})$ and $\psi(\boldsymbol{q})$ are the integration 'constants' which can be determined by the initial values $\delta(t_{\mathrm{mat}}, \boldsymbol{q})$ and $\dot{\delta}(t_{\mathrm{mat}}, \boldsymbol{q})$.

However, equation (362) can, unlike equation (357), also be derived from the General Theory of Relativity, and is, as a consequence, a *relativistic* equation: see Appendix E for a derivation. In this case, however, it is based on the gauge dependent quantity $\delta = \varepsilon_{(1)}/\varepsilon_{(0)}$. As a consequence, the second term of (361) is equal to the gauge mode

$$\delta_{\mathrm{gauge}}(\tau) = \psi\frac{\dot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} = -3H(t_{\mathrm{mat}})\psi\tau^{-1}, \tag{363}$$

as follows from (2a), (267b) and (268a). Therefore, the constant ψ in (363) and, hence, the function $\psi(\boldsymbol{q})$ in (359), should *not* be interpreted as an integration constant, but as a gauge function, which cannot be determined by imposing initial value conditions (see Appendix C). Since the solution (359) of equation (357) depends on the gauge function $\psi(\boldsymbol{q})$ it has no physical significance. Consequently, the standard equation (357) does *not* describe the evolution of density perturbations.

Again, we encounter the negative effect of the gauge function: up till now it was commonly accepted that for small-scale density perturbations (i.e. density perturbations with wave lengths much smaller than the particle horizon, Appendix D) the Newtonian theory suffices and gauge ambiguities do not occur and that the evolution of density perturbations in the Newtonian regime is described by the standard equation (357). The treatise presented in this article reveals, however, that in a fluid with an equation of state (264), the evolution of density perturbations is described by the *relativistic* equation (293), for small-scale as well as large-scale perturbations. In fact, equation (293) explains the formation of stars in the universe.

For pedagogical reasons only, we have calculated the 'formation of stars,' starting at $z = 1$, with the help of the standard equation (357) for which the conditions (308) hold true. The result is depicted in Figure 4. This figure should be compared with Figure 3, which is calculated with the help of the new equation (293) under the conditions (308). The standard Newtonian theory underestimates the internal gravitational field of a perturbation by a factor of $\frac{5}{3}$ with respect to our perturbation theory. Compare the gravitational fields $\frac{1}{2}\kappa\varepsilon_{(0)}$ in equation (357) and its relativistic counterpart (362) with $\frac{5}{6}\kappa\varepsilon_{(0)}$ in the new equation (293). As a result, the standard theory predicts a lower limit of $1.7\,\mathrm{M}_\odot$ for star formation, whereas our theory yields a lower limit of $0.8\,\mathrm{M}_\odot$ under the conditions (308). As a consequence, our Sun could not exist at all according to the standard theory. This fact can be considered as an 'experimental proof' that the standard Newtonian theory has indeed no physical significance.

Appendix A: Equations of State for the Energy Density and Pressure

We have used an equation of state for the pressure of the form $p = p(n, \varepsilon)$. In general, however, this equation of state is given in the form of two equations for the energy density ε and the pressure p which contain also the absolute temperature T:

$$\varepsilon = \varepsilon(n, T), \quad p = p(n, T). \tag{A1}$$

In principle it is possible to eliminate T from the two equations (A1) to get $p = p(n, \varepsilon)$, so that our choice of the form $p = p(n, \varepsilon)$ is justified. In practice, however, it may in general be difficult to eliminate the temperature T from the

equations (A1). However, this is not necessary, since the partial derivatives p_ε and p_n (78), the only quantities that are actually needed, can be found in an alternative way. From equations (A1) it follows

$$\mathrm{d}\varepsilon = \left(\frac{\partial\varepsilon}{\partial n}\right)_T \mathrm{d}n + \left(\frac{\partial\varepsilon}{\partial T}\right)_n \mathrm{d}T, \tag{A2a}$$

$$\mathrm{d}p = \left(\frac{\partial p}{\partial n}\right)_T \mathrm{d}n + \left(\frac{\partial p}{\partial T}\right)_n \mathrm{d}T. \tag{A2b}$$

From (A2b) it follows that the partial derivatives (78) are

$$p_n = \left(\frac{\partial p}{\partial n}\right)_T + \left(\frac{\partial p}{\partial T}\right)_n \left(\frac{\partial T}{\partial n}\right)_\varepsilon, \tag{A3a}$$

$$p_\varepsilon = \left(\frac{\partial p}{\partial T}\right)_n \left(\frac{\partial T}{\partial \varepsilon}\right)_n. \tag{A3b}$$

From (A2a) it follows

$$\left(\frac{\partial T}{\partial n}\right)_\varepsilon = -\left(\frac{\partial\varepsilon}{\partial n}\right)_T \left(\frac{\partial\varepsilon}{\partial T}\right)_n^{-1}, \tag{A4a}$$

$$\left(\frac{\partial T}{\partial \varepsilon}\right)_n = \left(\frac{\partial\varepsilon}{\partial T}\right)_n^{-1}. \tag{A4b}$$

Upon substituting the expressions (A4) into (A3), we find for the partial derivatives defined by (78)

$$p_n \equiv \left(\frac{\partial p}{\partial n}\right)_\varepsilon = \left(\frac{\partial p}{\partial n}\right)_T - \left(\frac{\partial p}{\partial T}\right)_n \left(\frac{\partial\varepsilon}{\partial n}\right)_T \left(\frac{\partial\varepsilon}{\partial T}\right)_n^{-1}, \tag{A5a}$$

$$p_\varepsilon \equiv \left(\frac{\partial p}{\partial \varepsilon}\right)_n = \left(\frac{\partial p}{\partial T}\right)_n \left(\frac{\partial\varepsilon}{\partial T}\right)_n^{-1}, \tag{A5b}$$

where ε and p are given by (A1). In order to calculate the second-order derivative p_{nn} replace p in (A5a) by p_n. For $p_{\varepsilon\varepsilon}$ replace p in (A5b) by p_ε. Finally, for $p_{\varepsilon n} \equiv p_{n\varepsilon}$, replace p in (A5a) by p_ε or, equivalently, replace p in (A5b) by p_n.

Appendix B: Derivation of the Manifestly Gauge-invariant Perturbation Equations

In this appendix we derive the perturbation equations (197) and the evolution equations (201).

1. Derivation of the Evolution Equation for the Entropy

With the help of equations (195a) and (195b) and equations (153)–(154) one may verify that

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\left(n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)}\varepsilon_{(1)}\right) = -3H\left(1 - \frac{n_{(0)}p_n}{\varepsilon_{(0)}(1+w)}\right)\left(n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)}\varepsilon_{(1)}\right). \tag{B1}$$

In view of (174) one may replace $\varepsilon_{(1)}$ and $n_{(1)}$ by $\varepsilon^{\mathrm{gi}}_{(1)}$ and $n^{\mathrm{gi}}_{(1)}$. Using (177) yields equation (197b) of the main text.

2. Derivation of the Evolution Equation for the Energy Density Perturbation

We will now derive equation (197a). To that end, we rewrite the system (195) and expression (196a) in the form

$$\dot\varepsilon_{(1)} + \alpha_{11}\varepsilon_{(1)} + \alpha_{12}n_{(1)} + \alpha_{13}\vartheta_{(1)} + \alpha_{14}\,{}^3\!R_{(1)\parallel} = 0, \tag{B2a}$$

$$\dot n_{(1)} + \alpha_{21}\varepsilon_{(1)} + \alpha_{22}n_{(1)} + \alpha_{23}\vartheta_{(1)} + \alpha_{24}\,{}^3\!R_{(1)\parallel} = 0, \tag{B2b}$$

$$\dot\vartheta_{(1)} + \alpha_{31}\varepsilon_{(1)} + \alpha_{32}n_{(1)} + \alpha_{33}\vartheta_{(1)} + \alpha_{34}\,{}^3\!R_{(1)\parallel} = 0, \tag{B2c}$$

$${}^3\!\dot R_{(1)\parallel} + \alpha_{41}\varepsilon_{(1)} + \alpha_{42}n_{(1)} + \alpha_{43}\vartheta_{(1)} + \alpha_{44}\,{}^3\!R_{(1)\parallel} = 0, \tag{B2d}$$

$$\varepsilon^{\mathrm{gi}}_{(1)} + \alpha_{51}\varepsilon_{(1)} + \alpha_{52}n_{(1)} + \alpha_{53}\vartheta_{(1)} + \alpha_{54}\,{}^3\!R_{(1)\parallel} = 0, \tag{B2e}$$

Table I: The coefficients α_{ij} figuring in the equations (B2).

$3H(1+p_\varepsilon)+\dfrac{\kappa\varepsilon_{(0)}(1+w)}{2H}$	$3Hp_n$	$\varepsilon_{(0)}(1+w)$	$\dfrac{\varepsilon_{(0)}(1+w)}{4H}$
$\dfrac{\kappa n_{(0)}}{2H}$	$3H$	$n_{(0)}$	$\dfrac{n_{(0)}}{4H}$
$\dfrac{p_\varepsilon}{\varepsilon_{(0)}(1+w)}\dfrac{\tilde{\nabla}^2}{a^2}$	$\dfrac{p_n}{\varepsilon_{(0)}(1+w)}\dfrac{\tilde{\nabla}^2}{a^2}$	$H(2-3\beta^2)$	0
$\dfrac{\kappa\,{}^3\!R_{(0)}}{3H}$	0	$-2\kappa\varepsilon_{(0)}(1+w)$	$2H+\dfrac{{}^3\!R_{(0)}}{6H}$
$\dfrac{-\,{}^3\!R_{(0)}}{{}^3\!R_{(0)}+3\kappa\varepsilon_{(0)}(1+w)}$	0	$\dfrac{6\varepsilon_{(0)}H(1+w)}{{}^3\!R_{(0)}+3\kappa\varepsilon_{(0)}(1+w)}$	$\dfrac{\frac{3}{2}\varepsilon_{(0)}(1+w)}{{}^3\!R_{(0)}+3\kappa\varepsilon_{(0)}(1+w)}$

where the coefficients $\alpha_{ij}(t)$ are given in Table I.

In calculating the coefficients a_1, a_2 and a_3, (198) in the main text, we use that the time derivative of the quotient w, defined by (153) is given by

$$\dot{w}=3H(1+w)(w-\beta^2), \tag{B3}$$

as follows from equation (154c) and the expression (153). Moreover, it is of convenience *not* to expand the function $\beta(t)$ defined by (153) since this will considerably complicate the expressions for the coefficients a_1, a_2 and a_3.

a. Step 1. We first eliminate the quantity ${}^3\!R_{(1)\parallel}$ from equations (B2). Differentiating equation (B2e) with respect to time and eliminating the time derivatives $\dot{\varepsilon}_{(1)}$, $\dot{n}_{(1)}$, $\dot{\vartheta}_{(1)}$ and ${}^3\!\dot{R}_{(1)\parallel}$ with the help of equations (B2a)–(B2d), we arrive at the equation

$$\dot{\varepsilon}^{\mathrm{gi}}_{(1)}+p_1\varepsilon_{(1)}+p_2n_{(1)}+p_3\vartheta_{(1)}+p_4\,{}^3\!R_{(1)\parallel}=0, \tag{B4}$$

where the coefficients $p_1(t),\ldots,p_4(t)$ are given by

$$p_i=\dot{\alpha}_{5i}-\alpha_{51}\alpha_{1i}-\alpha_{52}\alpha_{2i}-\alpha_{53}\alpha_{3i}-\alpha_{54}\alpha_{4i}. \tag{B5}$$

From equation (B4) it follows that

$${}^3\!R_{(1)\parallel}=-\frac{1}{p_4}\dot{\varepsilon}^{\mathrm{gi}}_{(1)}-\frac{p_1}{p_4}\varepsilon_{(1)}-\frac{p_2}{p_4}n_{(1)}-\frac{p_3}{p_4}\vartheta_{(1)}. \tag{B6}$$

In this way we have expressed the quantity ${}^3\!R_{(1)\parallel}$ as a linear combination of the quantities $\dot{\varepsilon}^{\mathrm{gi}}_{(1)}$, $\varepsilon_{(1)}$, $n_{(1)}$ and $\vartheta_{(1)}$. Upon replacing ${}^3\!R_{(1)\parallel}$ given by (B6), in equations (B2), we arrive at the system of equations

$$\dot{\varepsilon}_{(1)}+q_1\dot{\varepsilon}^{\mathrm{gi}}_{(1)}+\beta_{11}\varepsilon_{(1)}+\beta_{12}n_{(1)}+\beta_{13}\vartheta_{(1)}=0, \tag{B7a}$$
$$\dot{n}_{(1)}+q_2\dot{\varepsilon}^{\mathrm{gi}}_{(1)}+\beta_{21}\varepsilon_{(1)}+\beta_{22}n_{(1)}+\beta_{23}\vartheta_{(1)}=0, \tag{B7b}$$
$$\dot{\vartheta}_{(1)}+q_3\dot{\varepsilon}^{\mathrm{gi}}_{(1)}+\beta_{31}\varepsilon_{(1)}+\beta_{32}n_{(1)}+\beta_{33}\vartheta_{(1)}=0, \tag{B7c}$$
$${}^3\!\dot{R}_{(1)\parallel}+q_4\dot{\varepsilon}^{\mathrm{gi}}_{(1)}+\beta_{41}\varepsilon_{(1)}+\beta_{42}n_{(1)}+\beta_{43}\vartheta_{(1)}=0, \tag{B7d}$$
$$\varepsilon^{\mathrm{gi}}_{(1)}+q_5\dot{\varepsilon}^{\mathrm{gi}}_{(1)}+\beta_{51}\varepsilon_{(1)}+\beta_{52}n_{(1)}+\beta_{53}\vartheta_{(1)}=0, \tag{B7e}$$

where the coefficients $q_i(t)$ and $\beta_{ij}(t)$ are given by

$$q_i=-\frac{\alpha_{i4}}{p_4},\quad \beta_{ij}=\alpha_{ij}+q_ip_j. \tag{B8}$$

We now have achieved that the quantity ${}^3\!R_{(1)\parallel}$ occurs only in equation (B7d). Since we are not interested in the non-physical quantity ${}^3\!R_{(1)\parallel}$, we do not need this equation any more.

b. Step 2. We now proceed in the same way as in step 1: eliminating this time the quantity $\vartheta_{(1)}$ from the system of equations (B7). Differentiating equation (B7e) with respect to time and eliminating the time derivatives $\dot{\varepsilon}_{(1)}$, $\dot{n}_{(1)}$ and $\dot{\vartheta}_{(1)}$ with the help of equations (B7a)–(B7c), we arrive at

$$q_5\dddot{\varepsilon}^{\mathrm{gi}}_{(1)} + r\dot{\varepsilon}^{\mathrm{gi}}_{(1)} + s_1\varepsilon_{(1)} + s_2 n_{(1)} + s_3\vartheta_{(1)} = 0, \tag{B9}$$

where the coefficients $r(t)$ and $s_i(t)$ are given by

$$s_i = \dot{\beta}_{5i} - \beta_{51}\beta_{1i} - \beta_{52}\beta_{2i} - \beta_{53}\beta_{3i}, \tag{B10a}$$

$$r = 1 + \dot{q}_5 - \beta_{51}q_1 - \beta_{52}q_2 - \beta_{53}q_3. \tag{B10b}$$

From equation (B9) it follows that

$$\vartheta_{(1)} = -\frac{q_5}{s_3}\ddot{\varepsilon}^{\mathrm{gi}}_{(1)} - \frac{r}{s_3}\dot{\varepsilon}^{\mathrm{gi}}_{(1)} - \frac{s_1}{s_3}\varepsilon_{(1)} - \frac{s_2}{s_3}n_{(1)}. \tag{B11}$$

In this way we have expressed the quantity $\vartheta_{(1)}$ as a linear combination of the quantities $\ddot{\varepsilon}^{\mathrm{gi}}_{(1)}$, $\dot{\varepsilon}^{\mathrm{gi}}_{(1)}$, $\varepsilon_{(1)}$ and $n_{(1)}$. Upon replacing $\vartheta_{(1)}$ given by (B11) in equations (B7), we arrive at the system of equations

$$\dot{\varepsilon}_{(1)} - \beta_{13}\frac{q_5}{s_3}\ddot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(q_1 - \beta_{13}\frac{r}{s_3}\right)\dot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(\beta_{11} - \beta_{13}\frac{s_1}{s_3}\right)\varepsilon_{(1)} + \left(\beta_{12} - \beta_{13}\frac{s_2}{s_3}\right)n_{(1)} = 0, \tag{B12a}$$

$$\dot{n}_{(1)} - \beta_{23}\frac{q_5}{s_3}\ddot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(q_2 - \beta_{23}\frac{r}{s_3}\right)\dot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(\beta_{21} - \beta_{23}\frac{s_1}{s_3}\right)\varepsilon_{(1)} + \left(\beta_{22} - \beta_{23}\frac{s_2}{s_3}\right)n_{(1)} = 0, \tag{B12b}$$

$$\dot{\vartheta}_{(1)} - \beta_{33}\frac{q_5}{s_3}\ddot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(q_3 - \beta_{33}\frac{r}{s_3}\right)\dot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(\beta_{31} - \beta_{33}\frac{s_1}{s_3}\right)\varepsilon_{(1)} + \left(\beta_{32} - \beta_{33}\frac{s_2}{s_3}\right)n_{(1)} = 0, \tag{B12c}$$

$${}^3\dot{R}_{(1)\parallel} - \beta_{43}\frac{q_5}{s_3}\ddot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(q_4 - \beta_{43}\frac{r}{s_3}\right)\dot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(\beta_{41} - \beta_{43}\frac{s_1}{s_3}\right)\varepsilon_{(1)} + \left(\beta_{42} - \beta_{43}\frac{s_2}{s_3}\right)n_{(1)} = 0, \tag{B12d}$$

$$\varepsilon^{\mathrm{gi}}_{(1)} - \beta_{53}\frac{q_5}{s_3}\ddot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(q_5 - \beta_{53}\frac{r}{s_3}\right)\dot{\varepsilon}^{\mathrm{gi}}_{(1)} + \left(\beta_{51} - \beta_{53}\frac{s_1}{s_3}\right)\varepsilon_{(1)} + \left(\beta_{52} - \beta_{53}\frac{s_2}{s_3}\right)n_{(1)} = 0. \tag{B12e}$$

We have achieved now that the quantities $\vartheta_{(1)}$ and ${}^3R_{(1)\parallel}$ occur only in equations (B12c) and (B12d), so that these equations will not be needed anymore. We are left, in principle, with equations (B12a), (B12b) and (B12e) for the three unknown quantities $\varepsilon_{(1)}$, $n_{(1)}$ and $\varepsilon^{\mathrm{gi}}_{(1)}$, but we first proceed with all five equations.

c. Step 3. At first sight, the next steps would be to eliminate, successively, the quantities $\varepsilon_{(1)}$ and $n_{(1)}$ from equation (B12e) with the help of equations (B12a) and (B12b). We then would end up with a fourth-order differential equation for the unknown quantity $\varepsilon^{\mathrm{gi}}_{(1)}$.

However, it is possible to extract a second-order equation for the gauge-invariant energy density from the equations (B12). This will now be shown. Equation (B12e) can be rewritten

$$\ddot{\varepsilon}^{\mathrm{gi}}_{(1)} + a_1\dot{\varepsilon}^{\mathrm{gi}}_{(1)} + a_2\varepsilon^{\mathrm{gi}}_{(1)} = a_3\left(n_{(1)} + \frac{\beta_{51}s_3 - \beta_{53}s_1}{\beta_{52}s_3 - \beta_{53}s_2}\varepsilon_{(1)}\right), \tag{B13}$$

where the coefficients $a_1(t)$, $a_2(t)$ and $a_3(t)$ are given by

$$a_1 = -\frac{s_3}{\beta_{53}} + \frac{r}{q_5}, \quad a_2 = -\frac{s_3}{\beta_{53}q_5}, \quad a_3 = \frac{\beta_{52}s_3}{\beta_{53}q_5} - \frac{s_2}{q_5}. \tag{B14}$$

These are precisely the coefficients (198a)–(198c) of the main text. Furthermore, we find

$$\frac{\beta_{51}s_3 - \beta_{53}s_1}{\beta_{52}s_3 - \beta_{53}s_2} = -\frac{n_{(0)}}{\varepsilon_{(0)}(1+w)}. \tag{B15}$$

Hence,

$$n_{(1)} + \frac{\beta_{51}s_3 - \beta_{53}s_1}{\beta_{52}s_3 - \beta_{53}s_2}\varepsilon_{(1)} = n_{(1)} - \frac{n_{(0)}}{\varepsilon_{(0)}(1+w)}\varepsilon_{(1)}. \tag{B16}$$

With the help of this expression and (174) we can rewrite equation (B13) in the form (197a).

The derivation of the expressions (198) from (B14) and the proof of the equality (B15) is straightforward, but extremely complicated. We used Maple 13 [45] to perform this algebraic task [46].

3. Evolution Equations for the Contrast Functions

In this section we derive equations (201). We start off with equation (201b). From (177) and the definitions (200) it follows that

$$\sigma^{\mathrm{gi}}_{(1)} = n_{(0)}\left(\delta_n - \frac{\delta_\varepsilon}{1+w}\right). \tag{B17}$$

Differentiating this equation with respect to ct yields

$$a_4\sigma^{\mathrm{gi}}_{(1)} = \dot{n}_{(0)}\left(\delta_n - \frac{\delta_\varepsilon}{1+w}\right) + n_{(0)}\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta_n - \frac{\delta_\varepsilon}{1+w}\right), \tag{B18}$$

where we have used equation (197b). Using equation (154e) and the expression (B17) to eliminate $\sigma^{\mathrm{gi}}_{(1)}$, we arrive at equation (201b) of the main text.

Finally, we derive equation (201a). Upon substituting the expressions

$$\varepsilon^{\mathrm{gi}}_{(1)} = \varepsilon_{(0)}\delta_\varepsilon, \quad \dot{\varepsilon}^{\mathrm{gi}}_{(1)} = \dot{\varepsilon}_{(0)}\delta_\varepsilon + \varepsilon_{(0)}\dot{\delta}_\varepsilon, \quad \ddot{\varepsilon}^{\mathrm{gi}}_{(1)} = \ddot{\varepsilon}_{(0)}\delta_\varepsilon + 2\dot{\varepsilon}_{(0)}\dot{\delta}_\varepsilon + \varepsilon_{(0)}\ddot{\delta}_\varepsilon, \tag{B19}$$

into equation (197a), and dividing by $\varepsilon_{(0)}$, we find

$$b_1 = 2\frac{\dot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} + a_1, \quad b_2 = \frac{\ddot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} + a_1\frac{\dot{\varepsilon}_{(0)}}{\varepsilon_{(0)}} + a_2, \quad b_3 = a_3\frac{n_{(0)}}{\varepsilon_{(0)}}. \tag{B20}$$

where we have also used (B17). These are the coefficients (202) of the main text.

Appendix C: Gauge-invariance of the First-order Equations

If we go over from one synchronous system of reference with coordinates x to another synchronous system of reference with coordinates $\hat{x}$ given by expression (4), we have

$$\xi_{\mu;0} + \xi_{0;\mu} = 0, \tag{C1}$$

as follows from the transformation rule (11) and the conditions (27). From this equation we find, using (28), (30a), (30b) and (56) that $\xi^\mu(t,\boldsymbol{x})$ must be of the form

$$\xi^0 = \psi(\boldsymbol{x}), \quad \xi^i = \tilde{g}^{ik}\partial_k\psi(\boldsymbol{x})\int^{ct}\frac{\mathrm{d}\tau}{a^2(\tau)} + \chi^i(\boldsymbol{x}), \tag{C2}$$

where $\psi(\boldsymbol{x})$ and $\chi^i(\boldsymbol{x})$ are *arbitrary* functions —of the first-order— of the spatial coordinates $\boldsymbol{x}$. The fact that the gauge function ψ does not depend on the time coordinate $x^0 = ct$ anymore, as it did in general coordinates, see (6), is a consequence of the choice of synchronous coordinates for the original coordinates as well as for the transformed system of reference.

The energy density perturbation transforms according to (2a), where $\varepsilon_{(0)}$ is a solution of equation (154c). Similarly, the particle number density transforms according to (2b) where $n_{(0)}$ is a solution of equation (154e). Finally, as follows from (15), the fluid expansion scalar θ, (20c), transforms as

$$\hat{\theta}_{(1)}(t,\boldsymbol{x}) = \theta_{(1)}(t,\boldsymbol{x}) + \psi(\boldsymbol{x})\dot{\theta}_{(0)}(t), \tag{C3}$$

where $\theta_{(0)} = 3H$ is a solution of the set (154).

From (15) with $\sigma = p$, ε, or n and (77) we find for the transformation rule for the first-order perturbations to the pressure

$$\hat{p}_{(1)} = p_\varepsilon\hat{\varepsilon}_{(1)} + p_n\hat{n}_{(1)}. \tag{C4}$$

The transformation rule (16) with V^μ the four-velocity u^μ implies

$$\hat{u}^\mu_{(1)} = u^\mu_{(1)} - \xi^\mu{}_{,0}, \tag{C5}$$

where we have used that $u^\mu_{(0)} = \delta^\mu{}_0$, expression (53). From the transformation rule (C5) it follows that $u^\mu_{(1)}$ transforms under transformations (C2) between synchronous coordinates as

$$\hat{u}^0_{(1)}(t,\boldsymbol{x}) = u^0_{(1)}(t,\boldsymbol{x}) = 0, \tag{C6a}$$

$$\hat{u}^i_{(1)\|}(t,\boldsymbol{x}) = u^i_{(1)\|}(t,\boldsymbol{x}) - \frac{1}{a^2(t)}\tilde{g}^{ik}(\boldsymbol{x})\partial_k\psi(\boldsymbol{x}). \tag{C6b}$$

We want to determine the transformation rules for $\vartheta_{(1)}$ and ${}^3R_{(1)\|}$. Since the quantities ${}^3R$, (42), and ϑ, (47), are both non-scalars under general space-time transformations, the transformation rule (15) is not applicable to determine the transformation of their first-order perturbations under infinitesimal space-time transformations $x^\mu \to \hat{x}^\mu$, (4). Since $u^i_{(1)\|}$ satisfies equation (141e), and since $u^i_{(1)\|}$ transforms according to (C6), and since we know that $\hat{u}^i_{(1)\|}$ satisfies equation (141e) with hats, one may verify, using (C4), that

$$\hat{\vartheta}_{(1)}(t,\boldsymbol{x}) \equiv \vartheta_{(1)}(t,\boldsymbol{x}) - \frac{\tilde{\nabla}^2\psi(\boldsymbol{x})}{a^2(t)}, \tag{C7}$$

satisfies equation (161d) with hatted quantities. The quantity $\hat{\vartheta}_{(1)}$ is defined in analogy to $\vartheta_{(1)}$ in (74)

$$\hat{\vartheta}_{(1)} = (\hat{u}^k_{(1)\|})_{|k}. \tag{C8}$$

Apparently, $\vartheta_{(1)}$ transforms according to (C7) under arbitrary infinitesimal space-time transformations between synchronous coordinates. Similarly, one may verify that

$${}^3\hat{R}_{(1)\|}(t,\boldsymbol{x}) \equiv {}^3R_{(1)\|}(t,\boldsymbol{x}) + 4H(t)\left(\frac{\tilde{\nabla}^2\psi(\boldsymbol{x})}{a^2(t)} - \tfrac{1}{2}\,{}^3R_{(0)}(t)\psi(\boldsymbol{x})\right), \tag{C9}$$

satisfies equation (161a). Apparently, expression (C9) is the transformation rule for ${}^3R_{(1)\|}$ under arbitrary infinitesimal space-time transformations between synchronous coordinates. An alternative way to find the results (C7) and (C9) is to write $\hat{\vartheta}_{(1)} = \vartheta_{(1)} - f$ and ${}^3\hat{R}_{(1)\|} = {}^3R_{(1)\|} - g$, where f and g are unknown functions, to substitute, thereupon, $\hat{\vartheta}_{(1)}$ and ${}^3\hat{R}_{(1)\|}$ into equations (161d) and (161a), and to determine f and g such that the old equations (161d) and (161a) reappear. In fact, our method to define $\hat{\vartheta}_{(1)}$, (C7), and ${}^3\hat{R}_{(1)\|}$, (C9), is nothing but a shortcut to this procedure.

It may now easily be verified by substitution that if $\varepsilon_{(1)}$, $n_{(1)}$, $\theta_{(1)}$, $\vartheta_{(1)}$, and ${}^3R_{(1)\|}$ are solutions of the systems (152) and (195), then the quantities $\hat{\varepsilon}_{(1)}$, (2a), $\hat{n}_{(1)}$, (2b), $\hat{\theta}_{(1)}$, (C3), $\hat{\vartheta}_{(1)}$, (C7), and ${}^3\hat{R}_{(1)\|}$, (C9), are, for an arbitrary function $\psi(\boldsymbol{x})$, also solutions of these systems. In other words, the systems (152) and (195) are gauge-invariant under gauge transformations between synchronous coordinates. The solutions $\varepsilon_{(1)}$, $n_{(1)}$, $\theta_{(1)}$, $\vartheta_{(1)}$, and ${}^3R_{(1)\|}$, however, contain an arbitrary function $\psi(\boldsymbol{x})$ and are, therefore, gauge dependent.

Appendix D: Horizon Size after Decoupling of Matter and Radiation

As is well-known, a density perturbation can only grow if all its particles are in causal contact with each other, so that gravity can act in such a way that a density perturbation may eventually collapse. In this appendix we calculate the horizon size at time t_{mat} between the decoupling time t_{dec} and the present time t_{p}. The horizon size at time t is given by

$$d_{\mathrm{H}}(t) = ca(t)\int_0^t \frac{\mathrm{d}t'}{a(t')}. \tag{D1}$$

Using (269), we get

$$d_{\mathrm{H}}(t) = 3ct. \tag{D2}$$

For the horizon size at time t_{mat}, we find

$$d_{\mathrm{H}}(t_{\mathrm{mat}}) = 3ct_{\mathrm{mat}} = \frac{3ct_{\mathrm{p}}}{\big[z(t_{\mathrm{mat}})+1\big]^{3/2}} = \frac{1.260\times 10^7}{\big[z(t_{\mathrm{mat}})+1\big]^{3/2}}\ \mathrm{kpc}, \tag{D3}$$

where we have used (231), (232) and (269). At decoupling, $z(t_{\mathrm{dec}}) = 1091$, the horizon size is $d_{\mathrm{H}}(t_{\mathrm{dec}}) = 349\,\mathrm{kpc}$.

Appendix E: Derivation of the Relativistic Standard Equation for Density Perturbations

In this appendix equations (355) and (362) of the main text are derived, for a flat FLRW universe, ${}^3R_{(0)} = 0$, with vanishing cosmological constant, $\Lambda = 0$, using the background equations (154) and the linearized Einstein equations and conservation laws for scalar perturbations (195).

From (B3) it follows that w is constant if and only if $w = \beta^2$ for all times. Using (182) it is found for constant w that $p_n = 0$ and $p_\varepsilon = w$, i.e. the pressure does not depend on the particle number density. Consequently, in the derivation of equations (355) and (362) the equations (154e) for $n_{(0)}(t)$ and (195b) for $n_{(1)}(t, \boldsymbol{x})$ are not needed. In this case, the equation of state is given by

$$p = w\varepsilon. \tag{E1}$$

To derive the standard equations (355) and (362), it is required that $u^i_{(1)\parallel} = 0$, implying that

$$\vartheta_{(1)}(t, \boldsymbol{x}) = 0, \quad \psi(\boldsymbol{x}) = \psi, \tag{E2}$$

where we have used (218). The first of (E2) implies, using, equation (195c) that

$$\nabla^2 p_{(1)} = 0, \tag{E3}$$

i.e. pressure gradients should vanish in order to derive the *relativistic* standard equation (E6). Substituting $\varepsilon_{(1)} = \varepsilon_{(0)}\delta$ into equation (195a) and eliminating $\dot{\varepsilon}_{(0)}$ with the help of equation (154c), it is found that

$$\dot{\delta} + \frac{1+w}{2H}\left(\kappa\varepsilon_{(0)}\delta + \tfrac{1}{2}\,{}^3R_{(1)\parallel}\right) = 0. \tag{E4}$$

Differentiating equation (E4) with respect to $x^0 = ct$ and using equations (154) and (195d), yields

$$\ddot{\delta} + \tfrac{3}{2}(1+w)H\dot{\delta} - \tfrac{3}{4}(1+w)^2\kappa\varepsilon_{(0)}\delta - \tfrac{1}{8}(1+w)(1-3w)\,{}^3R_{(1)\parallel} = 0. \tag{E5}$$

Eliminating ${}^3R_{(1)\parallel}$ from equation (E5) with the help of equation (E4), yields the standard equation for large-scale perturbations in a flat FLRW universe:

$$\ddot{\delta} + 2H\dot{\delta} - \tfrac{1}{2}\kappa\varepsilon_{(0)}\delta(1+w)(1+3w) = 0. \tag{E6}$$

This equation has been derived by Weinberg [35], equation (15.10.57) and Peebles [9], equation (86.11). For $w = \frac{1}{3}$ (the radiation-dominated era) equation (355) is found, whereas for $w = 0$ (the era after decoupling of matter and radiation) equation (362) applies.

Using that the general solution of the background equations (154) for ${}^3R_{(0)} = 0$, $\Lambda = 0$ and constant w is given by

$$H(t) = \frac{2}{3(1+w)}(ct)^{-1} = H(t_0)\left(\frac{t}{t_0}\right)^{-1}, \tag{E7a}$$

$$\varepsilon_{(0)}(t) = \frac{4}{3\kappa(1+w)^2}(ct)^{-2} = \varepsilon_{(0)}(t_0)\left(\frac{t}{t_0}\right)^{-2}, \tag{E7b}$$

we find for the general solution of equation (E6), with $\tau \equiv t/t_0$,

$$\delta(\tau) = E_1\tau^{(2+6w)/(3+3w)} - 3(1+w)H(t_0)\psi\tau^{-1}, \tag{E8}$$

where E_1 is an arbitrary integration constant.

The only surviving gauge mode is, using (2a), (154c) and $\delta \equiv \varepsilon_{(1)}/\varepsilon_{(0)}$ [cf. (200)],

$$\varepsilon_{(1)\mathrm{gauge}}(t) = \psi\dot{\varepsilon}_{(0)}(t), \quad \delta_{\mathrm{gauge}}(t) = \frac{\varepsilon_{(1)\mathrm{gauge}}(t)}{\varepsilon_{(0)}(t)} = \psi\frac{\dot{\varepsilon}_{(0)}(t)}{\varepsilon_{(0)}(t)} = -3(1+w)H(t)\psi. \tag{E9}$$

It follows from (E7a) and (E9) that the second term in the right-hand side of (E8) is a gauge mode and ψ is the gauge constant. For $w = \frac{1}{3}$ we find (354) with gauge mode (356) and for $w = 0$ we get (361) with gauge mode (363). The solution (E8) is exactly equal to the result found by Peebles [9], §86, expression (86.12).

Finally, we note that in the derivation of (355) [i.e. (E6) with $w = \frac{1}{3}$] Peebles uses $\vartheta_{(1)} = 0$ (in his notation: $\theta = 0$), see (E2). In this case, Peebles' method yields a physical mode $\delta \propto \tau$ and a gauge mode $\delta \propto \tau^{-1}$. For $\vartheta_{(1)} \neq 0$ Peebles finds a physical mode $\delta \propto \tau^{1/2}$. However, in the treatise presented in this article both physical modes $\delta \propto \tau$ and $\delta \propto \tau^{1/2}$ [see (258)] follow from *one* second-order differential equation (245a). Moreover, in our treatise neither the pressure gradient $\boldsymbol{\nabla} p^{\mathrm{gi}}_{(1)}$, nor the velocity divergence $\vartheta_{(1)} \equiv (u^k_{(1)\|})_{|k}$ are neglected in our evolution equations (201).

Appendix F: Symbols and their Meaning

Table II: Symbols and their meaning of all quantities, except for those occurring in the appendices.

Symbol	Meaning	Reference Equation
$\boldsymbol{\nabla} f$	$(\partial_1 f, \partial_2 f, \partial_3 f)$	—
$(\tilde{\boldsymbol{\nabla}} f)^i$	$\tilde{g}^{ij}\partial_j f$	(125)–(128)
$\tilde{\boldsymbol{\nabla}}\cdot\boldsymbol{u}$	$u^k{}_{\vert k}$	(125), (126)
$(\tilde{\boldsymbol{\nabla}}\wedge\boldsymbol{u})_i$	$\epsilon_i{}^{jk}u_{j\vert k} = \epsilon_i{}^{jk}u_{j,k}$	(128)
$\tilde{\nabla}^2 f$	$\tilde{\boldsymbol{\nabla}}\cdot(\tilde{\boldsymbol{\nabla}} f) = \tilde{g}^{ij} f_{\vert i\vert j}$	(151)
∂_0	derivative with respect to $x^0 = ct$	—
∂_i	derivative with respect to x^i	—
hat: $\hat{}$	computed with respect to $\hat{x}$	(2)
dot: $\dot{}$	derivative with respect to $x^0 = ct$	(28)
prime: $'$	derivative with respect to τ	(253), (298)
tilde: $\tilde{}$	computed with respect to three-metric $\tilde{g}_{ij}$	(51b), (62)
super-index: $^{\mathrm{gi}}$	gauge-invariant	(3)
super-index: $^{\vert k}$	contravariant derivative with respect to x^k: $\zeta^{\vert k} = g^{kj}_{(0)}\zeta_{\vert j}$	(111a)
sub-index: $_{(0)}$	background quantity	(22)
sub-index: $_{(1)}$	perturbation of first-order	(22)
sub-index: $_{(2)}$	perturbation of second-order	(50)
sub-index: $_{;\lambda}$	covariant derivative with respect to x^λ	(7)
sub-index: $_{\vert k}$	covariant derivative with respect to x^k	(33)
sub-index: $_{,\mu}$	derivative with respect to x^μ	(29)
sub-index: $_{\parallel}$	longitudinal part of a vector or tensor	(110)
sub-index: $_{\perp}$	perpendicular part of vector or tensor	(110)
sub-index: $_{*}$	transverse and traceless part of a tensor	(110)
β	c^{-1} times speed of sound: $\sqrt{\dot{p}_{(0)}/\dot{\varepsilon}_{(0)}}$	(153)
$\Gamma^\lambda{}_{\mu\nu}$	connection coefficients	(29)
γ	arbitrary function	(113a)
δ_ε	energy contrast function	(200)
$\delta^\mu{}_\nu$	Kronecker delta	(53)
δ_n	particle number contrast function	(200)
δ_p	pressure contrast function	(206)
δ_T	(matter) temperature contrast function	(205), (286)
δ_{T_γ}	background radiation temperature contrast function	(205), (336)
ε	energy density	(20a)
e	energy per particle	(164)
$\epsilon_i{}^{jk}$	Levi-Civita tensor, $\epsilon_1{}^{23} = +1$	(128)
ζ, ϕ	potentials due to relativistic density perturbations	(111a), (207)
η	bookkeeping parameter equal to 1	(50)
θ	expansion scalar in four-space	(20c)
ϑ	expansion scalar in three-space	(47)
κ	$8\pi G/c^4$	(38)
$\varkappa_{ij}$	time derivative metric coefficients	(28)
Λ	cosmological constant	(36)
λ	wavelength, physical scale at time t_{p}	(248)
μ	thermodynamic potential	(167)
μ_{r}	reduced wave-number (radiation)	(254)
μ_{m}	reduced wave-number (matter)	(295)
ξ^μ	first-order space-time translation	(4)
π	arbitrary function	(113a)
ϖ	FLRW-coordinate	(52)
$\rho_{(1)}$	$\varepsilon^{\mathrm{gi}}_{(1)}/c^2$	(229)
σ	arbitrary scalar	(10a)
$\sigma^{\mathrm{gi}}_{(1)}$	abbreviation for $n^{\mathrm{gi}}_{(1)} - n_{(0)}\varepsilon^{\mathrm{gi}}_{(1)}/[\varepsilon_{(0)}(1+w)]$	(176), (177)
τ	dimensionless time	(251), (296)
ϕ, ζ	potentials due to relativistic density perturbations	(111a), (207)
φ	Newtonian potential	(225), (224)
χ^i	arbitrary three-vector	(C2)
ψ	first-order time translation	(2), (4), (6)
Ω	components in the Friedmann equation	(155)
ω	arbitrary scalar	(18)
$A^{\alpha\ldots\beta}{}_{\mu\ldots\nu}$	arbitrary tensor	(7)
$A_{\mu\nu}$	arbitrary rank two tensor	(10c)
a	scale factor or radius of universe	(51b)
a_{B}	black body constant	(236)
a_1, a_2, a_3, a_4	coefficients in perturbation equations	(198)
b_1, b_2, b_3	coefficients in perturbation equations	(202)

Table II: (continued)

c	speed of light	—
$\mathrm{d}s^2$	line element in four-space	(26)
E	energy within volume V	(167)
G	Newton's gravitation constant	—
$g_{\mu\nu}$	metric tensor	(11)
$\hat{g}_{\mu\nu}$	metric with respect to $\hat{x}^\mu$	(11)
$\tilde{g}_{ij}$	time-independent metric of three-space	(51b)
H	$c^{-1}\mathcal{H}$	(55)
$\mathcal{H}$	Hubble function: $\mathcal{H} = (\mathrm{d}a/\mathrm{d}t)/a$	(54)
h_{ij}	minus first-order perturbation of metric	(70)
k	$k = -1$ (open), 0 (flat), +1 (closed)	(52)
k_B	Boltzmann's constant	(264)
$\mathcal{L}_\xi$	Lie derivative with respect to ξ^μ	(7)
m	mass of particle of cosmological fluid	(213)
$\mathrm{M}_\odot$	solar mass, 1.98892×10^{30} kg	—
m_H	baryonic (proton) or CDM mass	(264)
N	number of particles within volume V	(167)
N^μ	particle density four-flow	(21)
n	particle number density	(20b)
p	pressure	(44)
p_ε, p_n	partial derivatives of pressure	(78)
$p_{nn}, p_{\varepsilon n}, p_{\varepsilon\varepsilon}$	partial derivatives of pressure	(199)
q	circular wave number: $2\pi/\lambda$	(248)
r	FLRW-coordinate	(52)
${}^3R$	Ricci scalar in three-space	(42)
$R_{\mu\nu}$	Ricci tensor in four-space	(31)
${}^3R_{ij}$	Ricci tensor in three-space	(34)
S	entropy within a volume V	(167)
s	entropy per particle $s \equiv S/N$	(164)
$s^\mathrm{gi}_{(1)}$	gauge-invariant entropy perturbation	(173)–(174)
T	absolute temperature	(167)
T_γ	photon temperature	(265)
$T^{\mu\nu}$	energy momentum tensor	(43)
$T^\mathrm{gi}_{(1)}$	gauge-invariant temperature perturbation	(190)
t	cosmological time	—
t_0	initial cosmological time	—
t_dec	decoupling time	(231)
t_eq	matter-radiation equality (onset of matter-dominated era)	(231)
t_mat	initial time between decoupling and the present	(270)
t_p	present cosmological time (13.7 Gyr)	(231)
t_rad	onset of radiation-dominated era	(240)
U^μ	cosmological four-velocity $U^\mu U_\mu = c^2$	(20)
u^μ	$c^{-1}U^\mu$	—
$\boldsymbol{U}$	spatial velocity	(212)
$\boldsymbol{u}$	$c^{-1}\boldsymbol{U}$	(124)
V^μ	arbitrary four-vector	(10b)
v_s	speed of sound	(273)
w	pressure divided by energy density	(153)
x	space-time point $x^\mu = (ct, \boldsymbol{x})$	—
$\hat{x}^\mu$	locally transformed coordinates	(4)
$\boldsymbol{x}$	spatial point $\boldsymbol{x} = (x^1, x^2, x^3)$	—
$z(t)$	redshift at time t	(232)

Acknowledgments

The authors should like to take the opportunity of thanking Marie-Louise Pennings MA (lecturer of English at the NLDA) for her assistance in discovering the intricacies of the English language.

[1] Carlos S. Frenk. Simulating the Formation of Cosmic Structure. *Phil. Trans. Roy. Soc. Lond.*, 300:1277, 2002. URL http://arxiv.org/abs/astro-ph/0208219.

[2] Volker Springel, Carlos S. Frenk, and Simon D. M. White. The large-scale structure of the universe. *Nature*, 440:1137, 2006. URL http://arxiv.org/abs/astro-ph/0604561.

[3] A. Loeb. Let there be Light: the Emergence of Structure out of the Dark Ages in the Early Universe, 2008. URL http://arxiv.org/abs/0804.2258.

[4] Thomas H. Greif, Jarrett L. Johnson, Ralf S. Klessen, and Volker Bromm. The First Galaxies: Assembly, Cooling and the Onset of Turbulence, 2008. URL http://arxiv.org/abs/0803.2237.
[5] T. M. Nieuwenhuizen, C. H. Gibson, and R. E. Schild. Gravitational Hydrodynamics of Large Scale Structure Formation. *ArXiv e-prints*, jun 2009. URL http://arXiv.org/abs/0906.5087.
[6] E. M. Lifshitz and I. M. Khalatnikov. Investigations in Relativistic Rosmology. *Adv. Phys.*, 12:185–249, 1963.
[7] P. J. Adams and V. Canuto. Exact Solution of the Lifshitz Equations Governing the Growth of Fluctuations in Cosmology. *Physical Review D*, 12(12):3793–3799, 1975.
[8] D. W. Olson. Density Perturbations in Cosmological Models. *Physical Review D*, 14(2):327–331, 1976.
[9] P. J. E. Peebles. *The Large-Scale Structure of the Universe*. Princeton Series in Physics. Princeton University Press, New Jersey, 1980.
[10] E. W. Kolb and M. S. Turner. *The Early Universe*. Frontiers in Physics. Addison-Wesley, Reading, MA, 1990.
[11] W. H. Press and E. T. Vishniac. Tenacious Myths about Cosmological Perturbations larger than the Horizon Size. *The Astrophysical Journal*, 239:1–11, July 1980.
[12] J. M. Bardeen. Gauge-invariant Cosmological Perturbations. *Phys. Rev.*, D22:1882–1905, 1980.
[13] V. F. Mukhanov and H. A. Feldman and R. H. Brandenberger. Theory of Cosmological Perturbations. *Phys. Rep.*, 215(5 & 6):203–333, 1992.
[14] D. N. Spergel et al. First year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *Astrophys. J. Suppl.*, 148:175, 2003. URL http://arxiv.org/abs/astro-ph/0302209.
[15] E. Komatsu et al. Five-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation, 2008. URL http://arxiv.org/abs/0803.0547.
[16] G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, N. Jarosik, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, M. Halpern, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results, 2008. URL http://arXiv.org/abs/0803.0732.
[17] J. Dunkley, E. Komatsu, M. R. Nolta, D. N. Spergel, D. Larson, G. Hinshaw, L. Page, C. L. Bennett, B. Gold, N. Jarosik, J. L. Weiland, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Likelihoods and Parameters from the WMAP Data. *ApJS*, 180:306–329, February 2009. doi: 10.1088/0067-0049/180/2/306. URL http://arXiv.org/abs/0803.0586.
[18] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ArXiv e-prints*, January 2010. URL http://arXiv.org/abs/1001.4538.
[19] Jarrett L. Johnson. Population III Star Clusters in the Reionized Universe, 2009. URL http://arXiv.org/abs/0911.1294.
[20] S. C. O. Glover, P. C. Clark, T. H. Greif, J. L. Johnson, V. Bromm, R. S. Klessen, and A. Stacy. Open Questions in the Study of Population III Star Formation, 2008. URL http://arXiv.org/abs/0808.0608.
[21] Michael L. Norman. Population III Star Formation and IMF, 2008. URL http://arXiv.org/abs/0801.4924.
[22] E. M. Lifshitz. On the Gravitational Stability of the Expanding Universe. *J. Phys.*, X(2):116–129, 1946.
[23] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields*, volume 2 of *Course of Theoretical Physics*. Pergamon Press, Oxford, 3 edition, 1975.
[24] S. W. Hawking. Perturbations of an Expanding Universe. *Astrophysical Journal*, 145(2):544–554, 1966.
[25] H. Kodama and M. Sasaki. Cosmological Perturbation Theory. *Progress of Theoretical Physics Supplement*, (78):1–166, 1984.
[26] G. F. R. Ellis and M. Bruni. Covariant and Gauge-invariant Approach to Cosmological Density Fluctuations. *Phys. Rev. D*, 40(6):1804–1818, 1989.
[27] G. F. R. Ellis, J. Hwang, and M. Bruni. Covariant and Gauge-independent Perfect-Fluid Robertson-Walker Perturbations. *Phys. Rev. D*, 40(6):1819–1826, 1989.
[28] George F. R. Ellis and Henk van Elst. Cosmological Models (Cargèse lectures 1998), 1998. URL http://arxiv.org/abs/gr-qc/9812046.
[29] V. Mukhanov. *Physical Foundations of Cosmology*. Cambridge University Press, 2005.
[30] E. Bertschinger. Cosmological dynamics. In M. Spiro R. Schaeffer, J. Silk and J. Zinn-Justin, editors, *Cosmologie et Structure à Grande Échelle*, pages 273–374. Elsevier Science, 1996. Les Houches, Session LX, 1993.
[31] Karim A. Malik and David Wands. Cosmological Perturbations, 2009. URL http://arXiv.org/abs/0809.4944.
[32] J. Hwang and H. Noh. Newtonian Limits of the Relativistic Cosmological Perturbations, 1997. URL http://arXiv.org/abs/astro-ph/9701137.
[33] J. Hwang and H. Noh. Why Newton's Gravity is practically reliable in the large-scale Cosmological Simulations, 2005. URL http://arXiv.org/abs/astro-ph/0507159.
[34] H. Noh and J. Hwang. Relativistic-Newtonian Correspondence of the Zero-pressure but weakly nonlinear Cosmology. *Class. Quant. Grav.*, 22:3181–3188, 2005. URL http://arxiv.org/abs/gr-qc/0412127.
[35] S. Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley & Sons, Inc. New York, 1972. ISBN 0-471-92567-5.
[36] J. W. York, jr. Covariant Decomposition of symmetric Tensors in the Theory of Gravitation. *Annales de l' Institut Henry Poincaré - Section A: Physique théorique*, XXI(4):319–332, 1974.
[37] J. M. Stewart. Perturbations of Friedmann-Robertson-Walker Cosmological Models. *Class. Quantum Grav.*, 7:1169–1180, 1990.

[38] The CDMS Collaboration and Z. Ahmed. Results from the Final Exposure of the CDMS II Experiment, 2009. URL http://arXiv.org/abs/0912.3592.

[39] T. Padmanabhan. *Structure Formation in the Universe*. Cambridge University Press, 1993.

[40] T. M. Nieuwenhuizen. Do Non-Relativistic Neutrinos constitute the Dark Matter? *Europhysics Letters*, 86:59001, June 2009. doi: 10.1209/0295-5075/86/59001. URL http://arXiv.org/abs/0812.4552.

[41] R Development Core Team. R: A Language and Environment for Statistical Computing, 2009. URL http://www.R-project.org.

[42] James B. Dent and Sourish Dutta. On the Dangers of using the Growth Equation on Large Scales, 2008. URL http://arXiv.org/abs/0808.2689.

[43] J. A. S. Lima, V. Zanchin, and R. Brandenberger. On the Newtonian Cosmology Equations with Pressure. *Mon. Not. Roy. Astron. Soc.*, 291:L1, 1997. URL http://arXiv.org/abs/astro-ph/9612166.

[44] J. A. Peacock. *Cosmological Physics*. Cambridge University Press, 1999. ISBN 0-521-42270-1.

[45] B. W. Char et al. *Maple V, Library Reference Manual*. Springer-Verlag, Berlin, Heidelberg, New York, 1992. ISBN 3-540-97592-6.

[46] One of the authors (PGM) has written a MAPLE program to check the final equations (197) and (201). This file will be send to the reader upon request.